

STIC-Biotech/ChemLib

156127

From: Seharaseyon, Jegatheesan  
Sent: Friday, June 10, 2005 2:32 PM  
To: STIC-Biotech/ChemLib  
Subject: RE: 10/008721

Hi,

Please search SEQ ID NO: 6 of 10/008721.  
Thanks

J. Seharaseyon  
Art Unit 1647  
Remsen 4C61  
Mailbox 4C70  
Phone: (571)-272-0892  
Fax: (571)-273-0892

CRFF

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: \_\_\_\_\_ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2005, 18:45:18 ; Search time 1435 Seconds  
(without alignments)  
4727.338 Million cell updates/sec

Title: US-10-008-721-6  
Perfect score: 140  
Sequence: 1 99gttcggaggaggatccc.....gagggtaaatagtgggcccag 140

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	140	100.0	140	6	AX477554 Sequence
2	140	100.0	140	6	AX477555 Sequence
3	140	100.0	140	6	AX504974 Sequence
4	140	100.0	140	6	AX504975 Sequence
5	140	100.0	771	6	AX477552 Sequence
6	140	100.0	771	6	AX477553 Sequence
7	140	100.0	771	6	AX504972 Sequence
8	140	100.0	771	6	AX504973 Sequence
9	140	100.0	955	6	AX477551 Sequence
10	140	100.0	955	6	AX504971 Sequence
11	140	100.0	955	9	AF315603 Homo sapi
12	140	100.0	31443	9	AC116341 Homo sapi
13	140	100.0	64466	9	AB036934 Homo sapi
14	140	100.0	90905	9	AY273797 Homo sapi
15	116.8	83.4	220823	2	AC150438 Saimiri s
16	53.4	38.1	207184	9	AC012361 Homo sapi
17	51.8	37.0	49477	9	BX284652 Human DNA
18	51.4	36.7	177883	2	AC046159 Homo sapi
19	51.4	36.7	207331	10	AC117235 Mus muscu

20	51.2	36.6	177675	9	AC008266 Homo sapi
21	51.2	36.6	197084	2	AC122937 Mus muscu
22	50.8	36.3	217726	10	AL928680 Mouse DNA
23	50.8	36.3	255615	2	AC098620 Rattus no
24	50.4	36.0	158063	9	AP001046 Homo sapi
25	50.4	36.0	191635	10	BX119996 Mouse DNA
26	50.4	36.0	248281	9	AC008737 Homo sapi
27	50.4	36.0	258174	2	AC079429 Mus muscu
28	50.4	36.0	340000	9	AP001751 Homo sapi
29	50.2	35.9	55239	2	AC068220 Homo sapi
30	50.2	35.9	231896	10	AC115116 Mus muscu
31	50	35.7	69941	2	AC027815 Homo sapi
32	49.8	35.6	113816	2	AC004358 Drosophil
33	49.8	35.6	147124	2	AC022768 Homo sapi
34	49.8	35.6	176052	10	AC140679 Mus muscu
35	49.8	35.6	183820	10	AL844533 Mouse DNA
36	49.8	35.6	195655	10	AC122824 Mus muscu
37	49.8	35.6	252055	2	AC106612 Rattus no
38	49.6	35.4	658	11	PM12G3B Penicilli
39	49.6	35.4	936	8	CNS01AYK AL113556 Botrytis
40	49.6	35.4	303091	2	AC084799 Mus muscu
41	49.4	35.3	3281	2	AC143602 Macaca mu
42	49.4	35.3	59909	2	AC090996 Homo sapi
43	49.4	35.3	110000	2	AC144302 1 of Continuation (2 of
44	49.4	35.3	127187	10	AL805908 Mouse DNA
45	49.4	35.3	156193	10	AL669907 Mouse DNA

ALIGNMENTS

RESULT 1  
AX477554  
LOCUS AX477554 140 bp DNA linear PAT 12-AUG-2002  
DEFINITION Sequence 6 from Patent WO0246433.  
ACCESSION AX477554  
VERSION AX477554.1 GI:22216734  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Saus,J.  
TITLE Tnf-inducible promoters and methods for using  
JOURNAL Patent: WO 0246433-A 6 13-JUN-2002;  
Saus, Juan (ES)  
FEATURES  
source Location/Qualifiers  
1..140  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN	Query Match	Best Local Similarity	Mismatches	Conservative	0;	Indels	0;	Gaps	0;
Qy	1	GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTTCAGACGCCGGGAGGGGACG	60						
Db	1	GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTTCAGACGCCGGGAGGGGACG	60						
Qy	61	GGCGGGGAGTACTGGGGGAGATGGAGACGAAAGGGGAGGAGGAGGAGGAGG	120						
Db	61	GGCGGGGAGTACTGGGGGAGATGGAGACGAAAGGGGAGGAGGAGGAGGAGG	120						
Qy	121	GAGGGTAAATAGTGGGCCAG	140						
Db	121	GAGGGTAAATAGTGGGCCAG	140						

RESULT 2  
AX477555/c

```
LOCUS AX477555 140 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 7 from Patent WO246433.
ACCESSION AX477555
VERSION AX477555.1 GI:22216735
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Tnf-inducible promoters and methods for using
JOURNAL Patent: WO 0246433-A 7 13-JUN-2002;
Saus, Juan (ES)
FEATURES
source Location/Qualifiers
1..140
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACCCGGGAGGGGACG 60
Db 140 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACCCGGGAGGGGACG 81
QY 61 GGGCGGGGAGTACTGGGGGAGAAATGGGAGGACGAAGGGAGGGGAAAGGACACAGGGGAGGG 120
Db 80 GGGCGGGGAGTACTGGGGGAGAAATGGGAGGACGAAGGGAGGGGAAAGGACACAGGGGAGGG 21
QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 20 GAGGGTAAATAGTGGGCCAG 1
RESULT 3
AX504974
LOCUS AX504974 140 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 6 from Patent WO246378.
ACCESSION AX504974
VERSION AX504974.1 GI:23386296
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Alternative pol k nucleotide and amino acid sequence and methods
JOURNAL Patent: WO 0246378-A 6 13-JUN-2002;
Saus, Juan (ES)
FEATURES
source Location/Qualifiers
1..140
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACCCGGGAGGGGACG 60
Db 140 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACCCGGGAGGGGACG 81
QY 61 GGGCGGGGAGTACTGGGGGAGAAATGGGAGGACGAAGGGAGGGGAAAGGACACAGGGGAGGG 120
Db 80 GGGCGGGGAGTACTGGGGGAGAAATGGGAGGACGAAGGGAGGGGAAAGGACACAGGGGAGGG 21
QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 20 GAGGGTAAATAGTGGGCCAG 1
RESULT 4
AX504975/c
LOCUS AX504975 140 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 7 from Patent WO246378.
ACCESSION AX504975
VERSION AX504975.1 GI:23386297
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Alternative pol k nucleotide and amino acid sequence and methods
JOURNAL Patent: WO 0246378-A 7 13-JUN-2002;
Saus, Juan (ES)
FEATURES
source Location/Qualifiers
1..140
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACCCGGGAGGGGACG 60
Db 140 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACCCGGGAGGGGACG 81
QY 61 GGGCGGGGAGTACTGGGGGAGAAATGGGAGGACGAAGGGAGGGGAAAGGACACAGGGGAGGG 120
Db 80 GGGCGGGGAGTACTGGGGGAGAAATGGGAGGACGAAGGGAGGGGAAAGGACACAGGGGAGGG 21
QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 20 GAGGGTAAATAGTGGGCCAG 1
RESULT 5
AX477552
LOCUS AX477552 771 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 4 from Patent WO246433.
ACCESSION AX477552
VERSION AX477552.1 GI:22216732
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Tnf-inducible promoters and methods for using
JOURNAL Patent: WO 0246433-A 4 13-JUN-2002;
Saus, Juan (ES)
FEATURES
source Location/Qualifiers
1..771
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 8.5e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 GGGTTCCGGAGGAGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 60
Db 577 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 636

QY 61 GGGCGGGAGTACTGGGGGAGATGGGAGACGAAAGGGGAGGAAAGGACAGGGGAGGG 120
Db 637 GGGCGGGAGTACTGGGGGAGATGGGAGGACGAAAGGGGAGGAAAGGACAGGGGAGGG 696

QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 697 GAGGGTAAATAGTGGGCCAG 716

RESULT 6
AX504972/c
LOCUS AX477553 771 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 5 from Patent WO0246378.
ACCESSION AX477553
VERSION AX477553.1 GI:22216733
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Tnf-inducible promoters and methods for using
JOURNAL Patent: WO 0246378-A 5 13-JUN-2002;
Saus, Juan (ES)
FEATURES
Location/Qualifiers
Source 1..771
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 8.5e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 60
Db 195 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 136

QY 61 GGGCGGGAGTACTGGGGGAGATGGGAGACGAAAGGGGAGGAAAGGACAGGGGAGGG 120
Db 135 GGGCGGGAGTACTGGGGGAGATGGGAGGACGAAAGGGGAGGAAAGGACAGGGGAGGG 76

QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 7
AX504972
LOCUS AX477551 771 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 4 from Patent WO0246378.
ACCESSION AX504972
VERSION AX504972.1 GI:23386294
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Alternative pol k nucleotide and amino acid sequence and methods
for using
JOURNAL Patent: WO 0246378-A 4 13-JUN-2002;
Saus, Juan (ES)
FEATURES
Location/Qualifiers
Source 1..771
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 8.5e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 60
Db 195 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 136

QY 61 GGGCGGGAGTACTGGGGGAGATGGGAGACGAAAGGGGAGGAAAGGACAGGGGAGGG 120
Db 135 GGGCGGGAGTACTGGGGGAGATGGGAGGACGAAAGGGGAGGAAAGGACAGGGGAGGG 76

QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 8
AX504973/c
LOCUS AX504973 771 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 5 from Patent WO0246378.
ACCESSION AX504973
VERSION AX504973.1 GI:23386295
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Alternative pol k nucleotide and amino acid sequence and methods
for using
JOURNAL Patent: WO 0246378-A 5 13-JUN-2002;
Saus, Juan (ES)
FEATURES
Location/Qualifiers
Source 1..771
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 8.5e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 60
Db 195 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 136

QY 61 GGGCGGGAGTACTGGGGGAGATGGGAGACGAAAGGGGAGGAAAGGACAGGGGAGGG 120
Db 135 GGGCGGGAGTACTGGGGGAGATGGGAGGACGAAAGGGGAGGAAAGGACAGGGGAGGG 76

QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 9
AX477551
LOCUS AX477551 955 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 3 from Patent WO0246433.
ACCESSION AX477551
VERSION AX477551.1 GI:22216731
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 8.5e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 60
Db 577 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 636

QY 61 GGGCGGGAGTACTGGGGGAGATGGGAGACGAAAGGGGAGGAAAGGACAGGGGAGGG 120
Db 637 GGGCGGGAGTACTGGGGGAGATGGGAGGACGAAAGGGGAGGAAAGGACAGGGGAGGG 696

QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 697 GAGGGTAAATAGTGGGCCAG 716

RESULT 8
AX504973/c
LOCUS AX504973 771 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 5 from Patent WO0246378.
ACCESSION AX504973
VERSION AX504973.1 GI:23386295
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Saus, J.
TITLE Alternative pol k nucleotide and amino acid sequence and methods
for using
JOURNAL Patent: WO 0246378-A 5 13-JUN-2002;
Saus, Juan (ES)
FEATURES
Location/Qualifiers
Source 1..771
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 140; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 8.5e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 60
Db 195 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTTCAGACGCCCGGGAGGGGACG 136

QY 61 GGGCGGGAGTACTGGGGGAGATGGGAGACGAAAGGGGAGGAAAGGACAGGGGAGGG 120
Db 135 GGGCGGGAGTACTGGGGGAGATGGGAGGACGAAAGGGGAGGAAAGGACAGGGGAGGG 76

QY 121 GAGGGTAAATAGTGGGCCAG 140
Db 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 9
AX477551
LOCUS AX477551 955 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 3 from Patent WO0246433.
ACCESSION AX477551
VERSION AX477551.1 GI:22216731
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Saus, J.  
TITLE Tnf-inducible promoters and methods for using

JOURNAL JOURNAL  
Saus, Juan (ES)  
Patent: WO 0246433-A 3 13-JUN-2002;

FEATURES  
source Location/Qualifiers

1..955  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 140; DB 6; Length 955;

Best Local Similarity 100.0%; Pred. No. 8e-19; Indels 0; Gaps 0;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60

Db 726 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 785

QY 61 GGGCGGGGAGTAGTGGGGGAGATGGGAGGACGAAGGGGAGGGAAGGACACGGGAGGG 120

Db 786 GGGCGGGGAGTAGTGGGGGAGATGGGAGGACGAAGGGGAGGGAAGGACACGGGAGGG 845

QY 121 GAGGGTAAATAGTGGGCCAG 140

Db 846 GAGGGTAAATAGTGGGCCAG 865

## RESULT 10

AX504971 AX504971 955 bp DNA linear PAT 27-SEP-2002

LOCUS Sequence 3 from Patent WO0246378.

DEFINITION AX504971

ACCESSION AX504971

VERSION AX504971.1 GI:23386293

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1

Saus, J.

Alternative pol k nucleotide and amino acid sequence and methods

for using

JOURNAL Patent: WO 0246378-A 3 13-JUN-2002;

Saus, Juan (ES)

## FEATURES

source

1..955

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 140; DB 6; Length 955;

Best Local Similarity 100.0%; Pred. No. 8e-19; Indels 0; Gaps 0;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60

Db 726 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 785

QY 61 GGGCGGGGAGTAGTGGGGGAGATGGGAGGACGAAGGGGAGGGAAGGACACGGGAGGG 120

Db 786 GGGCGGGGAGTAGTGGGGGAGATGGGAGGACGAAGGGGAGGGAAGGACACGGGAGGG 845

QY 121 GAGGGTAAATAGTGGGCCAG 140

Db 846 GAGGGTAAATAGTGGGCCAG 865

## RESULT 11

AF315603

AF315603 955 bp DNA linear PRI 20-SEP-2002

LOCUS Homo sapiens DNA polymerase kappa (POLK) and Goodpasture

autoantigen binding protein (COL4A3BP) genes, bidirectional

promoter and partial sequence.

ACCESSION AF315603

VERSION AF315603.1 GI:23208517

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 955)

Grano, F., Revert, F., Raya, A. and Saus, J.

A bidirectional promoter for the genes encoding DNA polymerase

kappa and Goodpasture autoantigen binding protein: Identification

of a novel pol kappa alternative spliced variant

Unpublished

JOURNAL 2 (bases 1 to 955)

Grano, F., Revert, F., Raya, A. and Saus, J.

Direct Submission

Submitted (23-OCT-2000) Patologia Molecular, Fundacion Valenciana

de Investigaciones Biomedicas, Amadeo de Saboya, 4, Valencia,

Valencia 46010, Spain

FEATURES

Location/Qualifiers

1..955

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="POLKGPBP18m-6C"

complement (<1..>865)

/genes="POLK"

complement (join (<1..443,567..>725))

/genes="POLK"

/product="DNA polymerase kappa"

complement (join (<1..443,567..>725))

/genes="POLK"

444..566

/number=1

<726..>955

/genes="COL4A3BP"

<726..>865

/notes="bidirectional"

mRNA

<866..>955

/genes="COL4A3BP"

/product="Goodpasture autoantigen binding protein"

<866..>955

/genes="COL4A3BP"

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60

Db 726 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 785

QY 61 GGGCGGGGAGTAGTGGGGGAGATGGGAGGACGAAGGGGAGGGAAGGACACGGGAGGG 120

Db 786 GGGCGGGGAGTAGTGGGGGAGATGGGAGGACGAAGGGGAGGGAAGGACACGGGAGGG 845

QY 121 GAGGGTAAATAGTGGGCCAG 140

Db 846 GAGGGTAAATAGTGGGCCAG 865

## RESULT 12

AC116341/c

LOCUS AC116341

DEFINITION Homo sapiens chromosome 5 clone RP11-144A5, complete sequence.

ACCESSION AC116341

VERSION AC116341.2 GI:21956545

KEYWORDS HTG.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 31443)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 31443)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-MAR-2002) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 REFERENCE 3 (bases 1 to 31443)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Jul 26, 2002 this sequence version replaced gi:19745021.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 100% of Sequence;  
 Estimated Total Number of Errors is 0.  
 NOTE: This insert is not the entire sequence of the clone (entire  
 sequence is 163kb). It is clipped at the overlaps with AC112183 and  
 AC026424. The number of bases overlapped with AC112183 is 15228  
 and with AC026424 is 3498.  
 FEATURES  
 Location/Qualifiers  
 1. .31443  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="RP11-144A5"  
 ORIGIN  
 Query Match 100.0%; Score 140; DB 9; Length 31443;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-19;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTTCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACGCGGAGGGGACG 60  
 DB 23487 GGGTTCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACGCGGAGGGGACG 23428  
 QY 61 GGGCGGGAGTGTGGGGGAGATGGGAGGACGAGGGGGGAGAGGACAGGGGAGGG 120  
 DB 23427 GGGCGGGAGTGTGGGGGAGATGGGAGGACGAGGGGGGAGAGGACAGGGGAGGG 23368  
 QY 121 GAGGGTAAATAGTGGGCCAG 140  
 DB 23367 GAGGGTAAATAGTGGGCCAG 23348  
 RESULT 13  
 AB036934/c  
 LOCUS 64466 bp DNA linear PRI 02-OCT-2001  
 DEFINITION Homo sapiens GPBP, DINB1 genes for Thr/Ser kinase, DINB1, partial  
 cds.  
 ACCESSION AB036934  
 VERSION AB036934.1 GI:14278874  
 KEYWORDS Thr/Ser kinase; DINB1.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Ogi, T., Yamamoto, Y. and Ohmori, H.  
 TITLE Homo sapiens genomic sequence, containing DINB1 & GPBP gene  
 JOURNAL Published Only in database (2001)  
 REFERENCE 2 (bases 1 to 64466)  
 AUTHORS Ogi, T., Yamamoto, Y. and Ohmori, H.

TITLE Direct Submission  
 JOURNAL Submitted (13-JAN-2000) Tomoo Ogi, Institute for Virus Research,  
 Kyoto University, Laboratory of Genetic Information Analysis;  
 Kawahara-machi, Syogoin, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan  
 (E-mail: togi@virus.kyoto-u.ac.jp, Tel: 81-75-751-4033 (ex.4033),  
 Fax: 81-75-751-3989)  
 FEATURES  
 Location/Qualifiers  
 1. .64466  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /map="5q13.1"  
 /clone="BAC CIT-C 784L6"  
 /notes="79.2cm"  
 mRNA complement(<3086. .3589)  
 gene complement(3086. .3181)  
 CDS /gene="GPBP"  
 complement(<3086. .3181)  
 /genes="GPBP"  
 /codon\_start=1  
 /product="Thr/Ser kinase"  
 /protein\_id="BAB58974.1"  
 /db\_xref="GI:14278875"  
 /translation="MSDNQNSGSEDEPESGPVERCGVLSK"  
 join(39145. .39279,44592. .44711,61466. .61618)  
 /genes="DINB1"  
 mRNA join(<39145. .39279,44592. .44711,61466. .>61618)  
 /genes="DINB1"  
 CDS join(39145. .39279,44592. .44711,61466. .>61618)  
 /genes="DINB1"  
 /notes="other name: POLK"  
 /codon\_start=1  
 /product="DINB1"  
 /protein\_id="BAB58975.1"  
 /db\_xref="GI:14278876"  
 /translation="MDSTKEKDSYKDDLLRMGFNDNKAGMEGLDKEKINKTKIMEAT  
 KGSFYFYNELKEKQVNRINENMQKQAITSQLRKAQLQVDRFAMELEQSRNLSNT  
 IVHIDMDAFYAAVEMRDNPELKDKPIAVGSMML"  
 ORIGIN  
 Query Match 100.0%; Score 140; DB 9; Length 64466;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-19;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTTCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACGCGGAGGGGACG 60  
 DB 3729 GGGTTCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACGCGGAGGGGACG 3670  
 QY 61 GGGCGGGAGTGTGGGGGAGATGGGAGGACGAGGGGGGAGGAGGAGGAGGG 120  
 DB 3669 GGGCGGGAGTGTGGGGGAGATGGGAGGACGAGGGGGGAGGAGGAGGAGGG 3610  
 QY 121 GAGGGTAAATAGTGGGCCAG 140  
 DB 3609 GAGGGTAAATAGTGGGCCAG 3590  
 RESULT 14  
 AY273797/c  
 LOCUS 90905 bp DNA linear PRI 22-APR-2003  
 DEFINITION Homo sapiens polymerase (DNA directed) kappa (POLK) gene, complete  
 cds.  
 ACCESSION AY273797  
 VERSION AY273797.1 GI:30039654  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 90905)  
 AUTHORS Rieder, M.J., Livingston, R.J., Daniels, M.R., Montoya, M.A.,  
 Chung, M.-W., Miyamoto, K.E., Nguyen, C.P., Nguyen, D.A., Poel, C.L.,





```

variation      11253
/gene="POLK"
/frequency="0.30"
/replace="n"
variation      11348
/gene="POLK"
/frequency="0.40"
/replace="c"
variation      11372
/gene="POLK"
/frequency="0.01"
/replace="c"
variation      11419
/gene="POLK"
/frequency="0.01"
/replace="c"
variation      11640
/gene="POLK"
/frequency="0.30"
/replace="a"
variation      11798
/gene="POLK"
/frequency="0.01"
/replace="g"
variation      11924
/gene="POLK"
/frequency="0.01"
/replace="c"
misc_feature   12008..12114
/gene="POLK"
/note="Region not scanned for variation"
repeat_region 12793..13209
/rpt_family="r2_type"
/rpt_type="dispersed"
variation      13335
/gene="POLK"
/frequency="0.01"
/replace="g"
variation      13443..13445
/gene="POLK"
/frequency="0.02"

Query Match      100.0%; Score 140; DB 9; Length 90905;
Best Local Similarity 100.0%; Pred. No. 2e-19;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGGCGTGTCCGTCAGACGCCGGAGGGGACG 60
|
|
|
Db 1984 GGGTTCCGGAGGAGGATCCCGAAGGCTCGGCGTGTCCGTCAGACGCCGGAGGGGACG 1925
|
|
|
QY 61 GGGCGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGGGAGGGGAAAGGACAGGGGAGGG 120
|
|
|
Db 1924 GGGCGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGGGAGGGGAAAGGACAGGGGAGGG 1865
|
|
|
QY 121 GAGGTAATAATAGTGGGCCAG 140
|
|
|
Db 1864 GAGGTAATAATAGTGGGCCAG 1845

RESULT 15
AC150438/c
LOCUS          220823 bp DNA linear HTG 20-JUL-2004
DEFINITION    Saimiri sciureus clone CH254-237A18, WORKING DRAFT SEQUENCE, 2
ordered pieces.
ACCESSION     AC150438
VERSION       AC150438.1 GI:50399918
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE        Saimiri sciureus (common squirrel monkey)
ORGANISM      Saimiri sciureus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae;
Saimiri.
REFERENCE     1 (bases 1 to 220823)

```

# AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,  
Peng,Z., Malinov,I. and Rubin,E.M.  
Direct Submission  
Unpublished  
2 (bases 1 to 220823)  
Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,  
Peng,Z., Malinov,I. and Rubin,E.M.  
Direct Submission  
Submitted (20-JUL-2004) Genome Sciences, Lawrence Berkeley National  
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

Sequence Produced by Berkeley PGA  
Web site: <http://pga.lbl.gov>  
Center Code: PGABERK  
Center Project Name: S070  
Bac Clone Name: CH254-237A18

This sequence has been compared to sequences of other species  
using Vista (<http://www-gsd.lbl.gov/Vista/>). The results can be  
viewed at:  
[http://pga.lbl.gov/cgi-bin/search\\_cvcgd?type=n&value=HMGCR](http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=HMGCR)

The order-orientation of the draft sequence was accomplished by  
using:  
Avid (<http://baboon.math.berkeley.edu/avid/>),  
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

## Summary Statistics:

Sequencing vector: Plasmid; pUC18  
Chemistry: Dye-terminator Big Dye  
Assembly Program: Phrap version 0.990329.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1 102039: contig of 102039 bp in length  
\* 102040 102139: gap of unknown length  
\* 102140 220823: contig of 118684 bp in length.

## FEATURES

source  
1..220823  
/organism="Saimiri sciureus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9521"  
/clone="CH254-237A18"

## ORIGIN

Query Match 83.4%; Score 116.8; DB 2; Length 220823;  
Best Local Similarity 94.3%; Pred. No. 7.4e-15;  
Matches 132; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGGCGTGTCCGTCAGACGCCGGAGGGGACG 60  
|  
|  
|  
Db 199443 GGGTCTGGAGGAGGATCCCGAAGGCTCGGCGTGTCCGTCAGACGCCGGAGGGGACG 199384  
|  
|  
|  
QY 61 GGGCGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGGGAGGGGAAAGGACAGGGGAGGG 120  
|  
|  
|  
Db 199383 GGGCGGGAGTAG-GAGGAGAAATGGGAGGACGAAGGGGAGGGGAAAGGACAGGGGAGGG 199325  
|  
|  
|  
QY 121 GAGGTAATAATAGTGGGCCAG 140  
|  
|  
|  
Db 199324 GAGGTAATAATAGTGGGCCAG 199305

Search completed: June 12, 2005, 20:43:54  
Job time : 1439 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2005, 18:45:18 ; Search time 279 Seconds  
(without alignments)  
2970.480 Million cell updates/sec

Title: US-10-008-721-6  
Perfect score: 140  
Sequence: 1 ggggtcggaggaggatccc.....gagggtaaatagtggggcag 140

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: , 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	140	6 AAD41044	Aad41044 Human COL
2	140	100.0	140	6 AAD41044	Abt03866 Human Goo
3	140	100.0	140	6 AAD41044	Abt03865 Human Goo
4	140	100.0	140	9 ACD13469	Accl13469 Human Pol
5	140	100.0	140	9 ACD13468	Accl13468 Human GPB
6	140	100.0	140	9 ADA97871	Ada97871 Human tum
7	140	100.0	140	9 ADA97870	Ada97870 Human tum
8	140	100.0	771	6 AAD41042	Aad41042 Human COL
9	140	100.0	771	6 AAD41042	Abt03863 Human Goo
10	140	100.0	771	6 AAD41042	Abt03864 Human Goo
11	140	100.0	771	9 ACD13467	Accl13467 Human Pol
12	140	100.0	771	9 ACD13466	Accl13466 Human GPB
13	140	100.0	771	9 ADA97869	Ada97869 Human tum
14	140	100.0	771	9 ADA97868	Ada97868 Human tum
15	140	100.0	955	6 AAD41041	Aad41041 Human 955
16	140	100.0	955	6 ACD13465	Accl13465 Human Goo
17	140	100.0	955	6 ACD13465	Accl13465 Human Pol
18	140	100.0	955	9 ADA97867	Ada97867 Human tum
19	48	34.3	26390	4 AAK65971	Aak65971 Human imm
20	47.6	34.0	408	13 ACN50226	Acn50226 Cotton no

21	47.2	33.7	1416	8	ABZ20967	Abz20967 Animal te
22	47	33.6	588	13	ACN54596	Acn54596 Cotton an
23	46.6	33.3	406	13	ACN59490	Acn59490 Cotton gy
24	46.6	33.3	20345	10	ADC86252	Adc86252 Human GPC
25	46.2	33.0	92219	13	ABD33589	Abd33589 Murine ca
26	45.6	32.6	629	13	ACN54594	Acn54594 Cotton an
27	45.2	32.3	1065	6	ABT09682	Abt09682 Human PAL
28	45	32.1	263	13	ACN61884	Acn61884 Cotton gy
29	45	32.1	1385	6	ABQ70117	Abq70117 Listeria
30	45	32.1	1385	6	ABQ68566	Abq68566 Listeria
31	45	32.1	23643	12	ADQ97331_5	Continuation (6 of
32	44.8	32.0	447	13	ACN55700	Acn55700 Cotton an
33	44.6	31.9	3952	4	AAL03258	Aal03258 Human rep
34	44.6	31.9	3955	4	AAL03257	Aal03257 Human rep
35	44.6	31.9	13591	5	AAF80047	Aaf80047 Nucleotid
36	44.4	31.7	31857	10	ADC86230	Adc86230 Human GPC
37	44.4	31.7	11750	13	ABD32653	Abd32653 Human can
38	44.2	31.6	2188	2	AAT77506	Aat77506 Human ova
39	44.2	31.6	3163	10	ADC87060	Adc87060 Human GPC
40	44	31.4	500	6	ABN15750	Abn15750 Human gen
41	44	31.4	500	13	ACN78840	Acn78840 Human GDM
42	43.8	31.3	987	6	ABT09669	Abt09669 Human PAL
43	43.4	31.0	410	4	AAI82055	Aai82055 Human pol
44	43.4	31.0	566	13	ACN59482	Acn59482 Cotton gy
45	43.4	31.0	600	6	ABQ52497	Abq52497 Oligonuci

ALIGNMENTS

RESULT 1  
AAD41044  
ID AAD41044 standard; DNA; 140 BP.  
AC AAD41044;  
XX  
DT 30-OCT-2002 (first entry)  
XX  
DE Human COL4A3BP DNA fragment.  
XX  
KW Human; tumour necrosis-factor; TNF; promoter; autoimmune disorder;  
KW cancer; therapy; db.  
XX  
OS Homo sapiens.  
XX  
PN WO200246433-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 07-DEC-2001; 2001WO-EP014412.  
XX  
PR 08-DEC-2000; 2000US-0254649P.  
XX  
PA (SAUS/) SAUS J.  
XX  
PI Saus J;  
XX  
DR WPI; 2002-519670/55.  
XX  
PT Novel tumor necrosis-factor inducible promoter useful for identifying  
PT candidate compounds for treating/preventing autoimmune disorders/cancer,  
PT or for identifying promoters that are regulated by tumor necrosis factor.  
PS Claim 1; Page 68; 95pp; English.  
XX  
CC The invention relates to a tumour necrosis-factor TNF inducible promoter.  
CC The invention is useful for identifying candidate TNF inducible promoters  
CC by aligning a test sequence consisting of a nucleic acid sequence with a  
CC comparison sequence selected from the invention, using a gap opening  
CC penalty of 50 and a gap extension penalty of 3 to define a test  
CC alignment, shuffling the nucleic sequence of the test sequence at least  
CC one hundred times, while maintaining its length and composition, to  
CC produce a series of randomised sequences, aligning the randomised

sequences with the comparison sequence using a gap opening penalty of 50 and a gap extension penalty of 3, to produce a series of randomised alignments, determining an average alignment quality of the randomised alignments, where the average alignment quality of the randomised alignments represent an alignment expected by chance, comparing the test alignment with the average alignment quality of the randomised alignments and identifying a test alignment with a probability value of less than 0.05 that the alignment is obtained by chance as a candidate TNF inducible promoter. The invention is useful for identifying candidate compounds for treating or preventing autoimmune disorders or cancer. The present sequence is human COL4A3BP DNA fragment

Sequence 140 BP; 31 A; 19 C; 77 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 140; DB 6; Length 140;  
Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACCGCGGGGAGG 60  
DB 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACCGCGGGGAGG 60

QY 61 GGGCGGGAGTACTGGGGGAGATGGAGGACCAAGGGGAGGAAAGGACAGGGGAGG 120  
DB 61 GGGCGGGAGTACTGGGGGAGATGGAGGACCAAGGGGAGGAAAGGACAGGGGAGG 120

QY 121 GAGGGTAAATAGTGGGCCAG 140  
DB 121 GAGGGTAAATAGTGGGCCAG 140

RESULT 2

ABT03866/c

ID ABT03866 standard; DNA; 140 BP.

AC ABT03866;

XX

DT 18-SEP-2002 (first entry)

XX

DE Human Goodpasture antigen binding protein gene fragment #7.

XX

KW Human; pol kappa 76; Goodpasture antigen binding protein; GPBP;  
KW\* chromosome 5q12-13; apoptosis; autoimmune disorder; cancer; cytostatic;  
KW immunosuppressive; gene; ds.

XX

OS Homo sapiens.

XX

PN WO200246378-A2.

XX

PD 13-JUN-2002.

XX

PF 07-DEC-2001; 2001WO-EP014409.

XX

PR 08-DEC-2000; 2000US-0254649P.

XX

PA (SAUS/) SAUS J.

XX

PI Saus J;

XX

DR WPI; 2002-537563/57.

XX

PT Novel isolated pol kappa76 polypeptide, a 76 kDa alternatively spliced variant of DNA polymerase kappa, useful as target for treating a patient with autoimmune disorder or cancer.

XX

PS Example; Page 63; 90pp; English.

XX

CC The present invention provides the protein and coding sequences of human DNA polymerase pol kappa 76. The gene is found on human chromosome 5q12-13, in a head-to-head arrangement with the Goodpasture antigen binding protein (GPBP). The detection of the coding sequence can be used for diagnosing an autoimmune condition and identifying cells undergoing apoptosis, and the sequences can be used in the treatment of autoimmune

CC diseases and cancer. The present sequence is a gene sequence described in the invention

XX

SQ Sequence 140 BP; 13 A; 77 C; 19 G; 31 T; 0 U; 0 Other;

Query Match 100.0%; Score 140; DB 6; Length 140;  
Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACCGCGGGGAGG 60  
DB 140 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCTCGCTCAGACCGCGGGGAGG 81

QY 61 GGGCGGGAGTACTGGGGGAGATGGAGGACCAAGGGGAGGAAAGGACAGGGGAGG 120  
DB 80 GGGCGGGAGTACTGGGGGAGATGGAGGACCAAGGGGAGGAAAGGACAGGGGAGG 21

QY 121 GAGGGTAAATAGTGGGCCAG 140  
DB 20 GAGGGTAAATAGTGGGCCAG 1

RESULT 3

ABT03865

ID ABT03865 standard; DNA; 140 BP.

AC ABT03865;

XX

DT 18-SEP-2002 (first entry)

XX

DE Human Goodpasture antigen binding protein gene fragment #6.

XX

KW Human; pol kappa 76; Goodpasture antigen binding protein; GPBP;  
KW\* chromosome 5q12-13; apoptosis; autoimmune disorder; cancer; cytostatic;  
KW immunosuppressive; gene; ds.

XX

OS Homo sapiens.

XX

PN WO200246378-A2.

XX

PD 13-JUN-2002.

XX

PF 07-DEC-2001; 2001WO-EP014409.

XX

PR 08-DEC-2000; 2000US-0254649P.

XX

PA (SAUS/) SAUS J.

XX

PI Saus J;

XX

DR WPI; 2002-537563/57.

XX

PT Novel isolated pol kappa76 polypeptide, a 76 kDa alternatively spliced variant of DNA polymerase kappa, useful as target for treating a patient with autoimmune disorder or cancer.

XX

PS Example; Page 63; 90pp; English.

XX

CC The present invention provides the protein and coding sequences of human DNA polymerase pol kappa 76. The gene is found on human chromosome 5q12-13, in a head-to-head arrangement with the Goodpasture antigen binding protein (GPBP). The detection of the coding sequence can be used for diagnosing an autoimmune condition and identifying cells undergoing apoptosis, and the sequences can be used in the treatment of autoimmune diseases and cancer. The present sequence is a gene sequence described in the invention

XX

SQ Sequence 140 BP; 31 A; 19 C; 77 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 140; DB 6; Length 140;  
Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTCTCGCTCAGACGCCGGAGGGGACG 60  
 Db 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTCTCGCTCAGACGCCGGAGGGGACG 60  
 Qy 61 GGGCGGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGAGGGGAAAGGACAGGGGGAGG 120  
 Db 61 GGGCGGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGAGGGGAAAGGACAGGGGGAGG 120  
 Qy 121 GAGGTTAAATAGTGGGCCAG 140  
 Db 121 GAGGTTAAATAGTGGGCCAG 140  
 RESULT 4  
 ACD13469/c  
 ID ACD13469 standard; DNA; 140 BP.  
 XX  
 AC ACD13469;  
 XT  
 DT 14-AUG-2003 (first entry)  
 XX  
 DE Human Pol kappa 76 intergene region/transcription start.  
 XX  
 KW Human; ds; Goodpasture antigen binding protein; GPBP; COL4A3BP;  
 KW collagen 4 alpha 3 binding protein; DNA polymerase kappa; Pol kappa;  
 KW Goodpasture disease; cutaneous lupus; polk76; bi-directional promoter;  
 KW autoimmune disease; cancer; antisense therapy; chromosome 5q12-13.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003027165-A1.  
 XX  
 PD 06-FEB-2003.  
 XX  
 XX 07-DEC-2001; 2001US-00010920.  
 XX  
 PF 08-DEC-2000; 2000US-0254649P.  
 XX  
 PA (SAUS/) SAUS J.  
 XX  
 PI Saus J;  
 XX  
 XX WPI; 2003-479531/45.  
 XX  
 XX New isolated DNA polymerase, pol kappa 76, useful in identifying  
 PT autoimmune disorders and in treating cancer and autoimmune disorders by  
 PT modifying its expression.  
 XX  
 PS Example; Page 21; 54pp; English.  
 XX  
 CC The invention relates to an isolated pol kappa (k) 76 polypeptide (an  
 CC alternatively spliced form of DNA polymerase kappa), appearing as  
 CC ABO07327 (encoded by the cDNA appearing as ACD13492). The gene for  
 CC POLKAPPA is located on chromosome 5q12-13 in a head-head arrangement with  
 CC the gene encoding Goodpasture antigen binding protein (GPBP or collagen 4  
 CC alpha 3 binding protein (COL4A3BP), associated with autoimmune diseases  
 CC such as Goodpasture's disease and cutaneous lupus) i.e. has a bi-  
 CC directional promoter. Also included are a recombinant expression vector  
 CC comprising the polk76 cDNA, a host cell transfected with the vector,  
 CC detecting (M1) polk76 (comprising providing a protein sample to be  
 CC screened, contacting the protein sample to be screened with an anti-  
 CC polk76 antibody and detecting the formation of an antibody-polypeptide  
 CC complexes, where the presence of the antibody-polypeptide complexes  
 CC indicates the presence of polk76), detecting (M2) the polk76 nucleic acid  
 CC in a sample (comprising contacting the sample with one or more polk76 PCR  
 CC primer, carrying out PCR to generate PCR products, and identifying the  
 CC polk76-specific PCR), detecting an autoimmune condition in a patient  
 CC (comprising providing a tissue or body fluid sample from the patient,  
 CC providing a control tissue or body fluid sample in which no autoimmune  
 CC condition is present, and detecting an increase in pol k76 RNA expression  
 CC in the tissue of body fluid samples compared to the control sample, where  
 CC the increase indicates the presence of an autoimmune condition) and  
 CC treating (M3) a patient with an autoimmune disorder or cancer by

CC modifying the expression or activity of pol k76 in the patient. Modifying  
 CC the expression or activity of polk76 or polk76 nucleic acid, such as by  
 CC increasing or decreasing their expression or activity using antibodies or  
 CC antisense therapy, is useful for treating an autoimmune disorder or  
 CC cancer. The present sequence is a PCR fragment representing part (or all)  
 CC of the intergenic region or bi-directional promoter of the polkappa/GPBP  
 CC genes  
 XX  
 SQ Sequence 140 BP; 13 A; 77 C; 19 G; 31 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 140; DB 9; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTCTCGCTCAGACGCCGGAGGGGACG 60  
 Db 140 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTCTCGCTCAGACGCCGGAGGGGACG 81  
 Qy 61 GGGCGGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGAGGGGAAAGGACAGGGGGAGG 120  
 Db 80 GGGCGGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGAGGGGAAAGGACAGGGGGAGG 21  
 Qy 121 GAGGTTAAATAGTGGGCCAG 140  
 Db 20 GAGGTTAAATAGTGGGCCAG 1  
 RESULT 5  
 ACD13468  
 ID ACD13468 standard; DNA; 140 BP.  
 XX  
 AC ACD13468;  
 XX  
 DT 14-AUG-2003 (first entry)  
 XX  
 DE Human GPBP/COL4A3BP intergene region/transcription start.  
 XX  
 KW Human; ds; Goodpasture antigen binding protein; GPBP; COL4A3BP;  
 KW collagen 4 alpha 3 binding protein; DNA polymerase kappa; Pol kappa;  
 KW Goodpasture disease; cutaneous lupus; polk76; bi-directional promoter;  
 KW autoimmune disease; cancer; antisense therapy; chromosome 5q12-13.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003027165-A1.  
 XX  
 PD 06-FEB-2003.  
 XX  
 PF 07-DEC-2001; 2001US-00010920.  
 XX  
 XX 08-DEC-2000; 2000US-0254649P.  
 XX  
 PA (SAUS/) SAUS J.  
 XX  
 PI Saus J;  
 XX  
 XX WPI; 2003-479531/45.  
 XX  
 PT New isolated DNA polymerase, pol kappa 76, useful in identifying  
 PT autoimmune disorders and in treating cancer and autoimmune disorders by  
 PT modifying its expression.  
 XX  
 PS Example; Page 21; 54pp; English.  
 XX  
 CC The invention relates to an isolated pol kappa (k) 76 polypeptide (an  
 CC alternatively spliced form of DNA polymerase kappa), appearing as  
 CC ABO07327 (encoded by the cDNA appearing as ACD13492). The gene for  
 CC POLKAPPA is located on chromosome 5q12-13 in a head-head arrangement with  
 CC the gene encoding Goodpasture antigen binding protein (GPBP or collagen 4  
 CC alpha 3 binding protein (COL4A3BP), associated with autoimmune diseases  
 CC such as Goodpasture's disease and cutaneous lupus) i.e. has a bi-  
 CC directional promoter. Also included are a recombinant expression vector  
 CC comprising the polk76 cDNA, a host cell transfected with the vector,  
 CC detecting (M1) polk76 (comprising providing a protein sample to be  
 CC screened, contacting the protein sample to be screened with an anti-  
 CC polk76 antibody and detecting the formation of an antibody-polypeptide  
 CC complexes, where the presence of the antibody-polypeptide complexes  
 CC indicates the presence of polk76), detecting (M2) the polk76 nucleic acid  
 CC in a sample (comprising contacting the sample with one or more polk76 PCR  
 CC primer, carrying out PCR to generate PCR products, and identifying the  
 CC polk76-specific PCR), detecting an autoimmune condition in a patient  
 CC (comprising providing a tissue or body fluid sample from the patient,  
 CC providing a control tissue or body fluid sample in which no autoimmune  
 CC condition is present, and detecting an increase in pol k76 RNA expression  
 CC in the tissue of body fluid samples compared to the control sample, where  
 CC the increase indicates the presence of an autoimmune condition) and  
 CC treating (M3) a patient with an autoimmune disorder or cancer by

CC detecting (M1) polk76 (comprising providing a protein sample to be  
 CC screened, contacting the protein sample to be screened with an anti-  
 CC polk76 antibody and detecting the formation of an antibody- polypeptide  
 CC complexes, where the presence of the antibody-polypeptide complexes  
 CC indicates the presence of polk76), detecting (M2) the polk76 nucleic acid  
 CC in a sample (comprising contacting the sample with one or more polk76 PCR  
 CC primer, carrying out PCR to generate PCR products, and identifying the  
 CC polk76-specific PCR), detecting an autoimmune condition in a patient  
 CC (comprising providing a tissue or body fluid sample from the patient,  
 CC providing a control tissue or body fluid sample in which no autoimmune  
 CC condition is present, and detecting an increase in pol k76 RNA expression  
 CC in the tissue of body fluid samples compared to the control sample, where  
 CC the increase indicates the presence of an autoimmune condition) and  
 CC treating (M3) a patient with an autoimmune disorder or cancer by  
 CC modifying the expression or activity of pol k76 in the patient. Modifying  
 CC the expression or activity of polk76 or polk76 nucleic acid, such as by  
 CC increasing or decreasing their expression or activity using antibodies or  
 CC antisense therapy, is useful for treating an autoimmune disorder or  
 CC cancer. The present sequence is a PCR fragment representing part (or all)  
 CC of the intergenic region or bi-directional promoter of the polkappa/GPBP  
 CC genes  
 XX  
 SQ Sequence 140 BP; 31 A; 19 C; 77 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 140; DB 9; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTGCGGTTCAGACGCCGGAGGGGACG 60  
 Db 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTGCGGTTCAGACGCCGGAGGGGACG 60  
 QY 61 GGGCGGGGAGTAGTGGGGAGGATGGAGACGAAGGGGGGGAAGGACAGCGGAGGG 120  
 Db 61 GGGCGGGGAGTAGTGGGGAGGATGGAGACGAAGGGGGGGAAGGACAGCGGAGGG 120  
 QY 121 GAGGGTAAATAGTGGGCCAG 140  
 Db 121 GAGGGTAAATAGTGGGCCAG 140

RESULT 6  
 ADA97871/c  
 ID ADA97871 standard; DNA; 140 BP.  
 AC ADA97871;  
 XX  
 XX 20-NOV-2003 (first entry)  
 XX Human tumour necrosis factor (TNF) inducible promoter #5.  
 XX  
 KW Human; tumour necrosis factor inducible promoter; TNF;  
 KW autoimmune disorder; cancer; ds; promoter; immunosuppressive; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX US2003082745-A1.  
 PN 01-MAY-2003.  
 PD 07-DEC-2001; 2001US-00008721.  
 PF 08-DEC-2000; 2000US-0254649P.  
 XX (SAUS/) SAUS J.  
 PA Saus J;  
 PI WPI; 2003-606062/57.  
 DR  
 XX  
 XX New tumor necrosis factor inducible promoters, useful for identifying  
 PT promoters that are regulated by tumor necrosis factor, or for identifying  
 PT candidate compounds for treating or preventing autoimmune disorders or

PT cancer.  
 XX  
 PS Claim 1; Fig 1; 57pp; English.  
 XX  
 CC The invention relates to a tumour necrosis factor (TNF) inducible  
 CC promoter. Also disclosed are an expression vector comprising one or more  
 CC tumour necrosis factor inducible promoters and a recombinant host cell  
 CC transfected with one or more expression vectors. The TNF inducible  
 CC promoters, expression vectors and host cells are useful for identifying  
 CC promoters that are regulated by tumour necrosis factor or for identifying  
 CC candidate compounds for treating or preventing autoimmune disorders or  
 CC cancer. This sequence represents a tumour necrosis factor inducible  
 CC promoter of the invention.  
 XX  
 SQ Sequence 140 BP; 13 A; 77 C; 19 G; 31 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 140; DB 9; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTGCGGTTCAGACGCCGGAGGGGACG 60  
 Db 140 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTGCGGTTCAGACGCCGGAGGGGACG 81  
 QY 61 GGGCGGGGAGTAGTGGGGAGGATGGAGACGAAGGGGGGGAAGGACAGCGGAGGG 120  
 Db 80 GGGCGGGGAGTAGTGGGGAGGATGGAGACGAAGGGGGGGAAGGACAGCGGAGGG 21  
 QY 121 GAGGGTAAATAGTGGGCCAG 140  
 Db 20 GAGGGTAAATAGTGGGCCAG 1  
 RESULT 7  
 ADA97870  
 ID ADA97870 standard; DNA; 140 BP.  
 AC ADA97870;  
 XX  
 XX 20-NOV-2003 (first entry)  
 XX Human tumour necrosis factor (TNF) inducible promoter #4.  
 XX  
 KW Human; tumour necrosis factor inducible promoter; TNF;  
 KW autoimmune disorder; cancer; ds; promoter; immunosuppressive; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX US2003082745-A1.  
 PN 01-MAY-2003.  
 PD 07-DEC-2001; 2001US-00008721.  
 PF 08-DEC-2000; 2000US-0254649P.  
 XX (SAUS/) SAUS J.  
 PA Saus J;  
 PI WPI; 2003-606062/57.  
 DR  
 XX  
 XX New tumor necrosis factor inducible promoters, useful for identifying  
 PT promoters that are regulated by tumor necrosis factor, or for identifying  
 PT candidate compounds for treating or preventing autoimmune disorders or  
 PT cancer.  
 XX  
 PS Claim 1; Fig 1; 57pp; English.  
 XX  
 CC The invention relates to a tumour necrosis factor (TNF) inducible  
 CC promoter. Also disclosed are an expression vector comprising one or more  
 CC tumour necrosis factor inducible promoters and a recombinant host cell  
 CC transfected with one or more expression vectors. The TNF inducible

CC promoters, expression vectors and host cells are useful for identifying  
 CC promoters that are regulated by tumour necrosis factor or for identifying  
 CC candidate compounds for treating or preventing autoimmune disorders or  
 CC cancer. This sequence represents a tumour necrosis factor inducible  
 CC promoter of the invention.

SQ Sequence 140 BP; 31 A; 19 C; 77 G; 13 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 140; DB 9; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTTCGGGAGGAGATCCGAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGAGCG 60  
 DB 1 GGGTTCGGGAGGAGATCCGAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGAGCG 60  
 QY 61 GGGCGGGGAGTGTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGG 120  
 DB 61 GGGCGGGGAGTGTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGG 120  
 QY 121 GAGGTTAAATAGTGGGCCAG 140  
 DB 121 GAGGTTAAATAGTGGGCCAG 140

RESULT 8  
 AAD41042  
 ID AAD41042 standard; DNA; 771 BP.  
 AC AAD41042;  
 XX  
 DT 30-OCT-2002 (first entry)  
 XX  
 DE Human COL4A3BP DNA.  
 XX  
 XX Human; tumour necrosis-factor; TNF; promoter; autoimmune disorder;  
 KW cancer; therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200246433-A2.  
 XX  
 PD 13-JUN-2002;  
 XX  
 PF 07-DEC-2001; 2001WO-EP014412.  
 XX  
 PR 08-DEC-2000; 2000US-0254649P.  
 XX  
 PA (SAUS/) SAUS J.  
 XX  
 PI Saus J;  
 XX  
 DR WPI; 2002-519670/55.  
 XX  
 XX Novel tumor necrosis-factor inducible promoter useful for identifying  
 PT candidate compounds for treating/preventing autoimmune disorders/cancer,  
 PT or for identifying promoters that are regulated by tumor necrosis factor.  
 XX  
 PS Claim 1; Page 67; 95pp; English.  
 XX  
 CC The invention relates to a tumour necrosis-factor TNF inducible promoter.  
 CC The invention is useful for identifying candidate TNF inducible promoters  
 CC by aligning a test sequence consisting of a nucleic acid sequence with a  
 CC comparison sequence selected from the invention, using a gap opening  
 CC penalty of 50 and a gap extension penalty of 3 to define a test  
 CC alignment, shuffling the nucleic sequence of the test sequence at least  
 CC one hundred times, while maintaining its length and composition, to  
 CC produce a series of randomised sequences, aligning the randomised  
 CC sequences with the comparison sequence using a gap opening penalty of 50  
 CC and a gap extension penalty of 3, to produce a series of randomised  
 CC alignments, determining an average alignment quality of the randomised  
 CC alignments, where the average alignment quality of the randomised  
 CC alignments represent an alignment expected by chance, comparing the test

CC alignment with the average alignment quality of the randomised alignments  
 CC and identifying a test alignment with a probability value of less than  
 CC 0.05 that the alignment is obtained by chance as a candidate TNF  
 CC inducible promoter. The invention is useful for identifying candidate  
 CC compounds for treating or preventing autoimmune disorders or cancer. The  
 CC present sequence is human COL4A3BP DNA

SQ Sequence 771 BP; 186 A; 209 C; 253 G; 123 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 140; DB 6; Length 771;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTTCGGGAGGAGATCCGAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGAGCG 60  
 DB 577 GGGTTCGGGAGGAGATCCGAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGAGCG 636  
 QY 61 GGGCGGGGAGTGTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGG 120  
 DB 637 GGGCGGGGAGTGTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGG 696  
 QY 121 GAGGTTAAATAGTGGGCCAG 140  
 DB 697 GAGGTTAAATAGTGGGCCAG 716

RESULT 9  
 ABT03863  
 ID ABT03863 standard; DNA; 771 BP.  
 XX  
 AC ABT03863;  
 XX  
 DT 18-SEP-2002 (first entry)  
 XX  
 DE Human Goodpasture antigen binding protein gene fragment #4.  
 XX  
 XX Human; pol kappa 76; Goodpasture antigen binding protein; GPBP;  
 KW chromosome 5q12-13; apoptosis; autoimmune disorder; cancer; cytostatic;  
 KW immunosuppressive; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200246378-A2.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 07-DEC-2001; 2001WO-EP014409.  
 XX  
 PR 08-DEC-2000; 2000US-0254649P.  
 XX  
 PA (SAUS/) SAUS J.  
 XX  
 PI Saus J;  
 XX  
 DR WPI; 2002-537563/57.  
 XX  
 XX Novel isolated pol kappa76 polypeptide, a 76 kDa alternatively spliced  
 PT variant of DNA polymerase kappa, useful as target for treating a patient  
 PT with autoimmune disorder or cancer.  
 XX  
 PS Example; Page 62; 90pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of human  
 CC DNA polymerase pol kappa 76. The gene is found on human chromosome 5q12-  
 CC 13, in a head-to-head arrangement with the Goodpasture antigen binding  
 CC protein (GPBP). The detection of the coding sequence can be used for  
 CC diagnosing an autoimmune condition and identifying cells undergoing  
 CC apoptosis, and the sequences can be used in the treatment of autoimmune  
 CC diseases and cancer. The present sequence is a gene sequence described in  
 CC the invention  
 XX  
 SQ Sequence 771 BP; 186 A; 209 C; 253 G; 123 T; 0 U; 0 Other;

Query Match 100.0%; Score 140; DB 6; Length 771;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTTCGGGAGGAGATCCGAAAGGCTCGCGTGTGCGCTCAGACGCCGGAGGGGAGCG 60  
 |||||  
 Db 577 GGTTTCGGGAGGAGATCCGAAAGGCTCGCGTGTGCGCTCAGACGCCGGAGGGGAGCG 636  
 |||||

QY 61 GGGCGGGAGTACTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGACAGGGGAGGG 120  
 |||||  
 Db 637 GGGCGGGAGTACTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGGAGGG 696  
 |||||

QY 121 GAGGGTAAATAGTGGGCCAG 140  
 |||||  
 Db 697 GAGGGTAAATAGTGGGCCAG 716  
 |||||

RESULT 10  
 ABT03864/c  
 ID ABT03864 standard; DNA; 771 BP.  
 XX  
 AC ABT03864;  
 XX  
 DT 18-SEP-2002 (first entry)  
 XX  
 DE Human Goodpasture antigen binding protein gene fragment #5.  
 XX  
 KW Human; pol kappa 76; Goodpasture antigen binding protein; GPBP;  
 KW chromosome 5q12-13; apoptosis; autoimmune disorder; cancer; cytostatic;  
 KW immunosuppressive; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200246378-A2.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 07-DEC-2001; 2001WO-EP014409.  
 XX  
 PR 08-DEC-2000; 2000US-0254649P.  
 XX  
 PA (SAUS/) SAUS J.  
 XX  
 PI Saus J;  
 XX  
 PS Example; Page 63; 90pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of human  
 CC DNA polymerase pol kappa 76. The gene is found on human chromosome 5q12-  
 CC 13, in a head-to-head arrangement with the Goodpasture antigen binding  
 CC protein (GPBP). The detection of the coding sequence can be used for  
 CC diagnosing an autoimmune condition and identifying cells undergoing  
 CC apoptosis, and the sequences can be used in the treatment of autoimmune  
 CC diseases and cancer. The present sequence is a gene sequence described in  
 CC the invention  
 XX  
 SQ Sequence 771 BP; 123 A; 253 C; 209 G; 186 T; 0 U; 0 Other;

Query Match 100.0%; Score 140; DB 6; Length 771;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTTCGGGAGGAGATCCGAAAGGCTCGCGTGTGCGCTCAGACGCCGGAGGGGAGCG 60  
 |||||  
 Db 195 GGTTTCGGGAGGAGGATCCGAAAGGCTCGCGTGTGCGCTCAGACGCCGGAGGGGAGCG 136  
 |||||

QY 61 GGGCGGGAGTACTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGGAGGG 120  
 |||||

Db 135 GGGCGGGAGTACTGGGGAGATGGGAGGACGAGGAGGGGAGGAGGAGGAGGG 76  
 |||||

QY 121 GAGGGTAAATAGTGGGCCAG 140  
 |||||

Db 75 GAGGGTAAATAGTGGGCCAG 56  
 |||||

RESULT 11  
 ACD13467/c  
 ID ACD13467 standard; DNA; 771 BP.  
 XX  
 AC ACD13467;  
 XX  
 DT 14-AUG-2003 (first entry)  
 XX  
 DE Human Pol kappa 76 771bp PCR fragment.  
 XX  
 KW Human; ds; Goodpasture antigen binding protein; GPBP; COL4A3BP;  
 KW collagen 4 alpha 3 binding protein; DNA polymerase kappa; Pol kappa;  
 KW Goodpasture disease; cutaneous lupus; polk76; bi-directional promoter;  
 KW autoimmune disease; cancer; antisense therapy; chromosome 5q12-13.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003027165-A1.  
 XX  
 PD 06-FEB-2003.  
 XX  
 PF 07-DEC-2001; 2001US-00010920.  
 XX  
 PR 08-DEC-2000; 2000US-0254649P.  
 XX  
 PA (SAUS/) SAUS J.  
 XX  
 PI Saus J;  
 XX  
 PS WPI; 2003-479531/45.  
 XX  
 CC New isolated DNA polymerase, pol kappa 76, useful in identifying  
 CC autoimmune disorders and in treating cancer and autoimmune disorders by  
 CC modifying its expression.  
 XX  
 PS Example; Page 21; 54pp; English.  
 XX  
 CC The invention relates to an isolated pol kappa (k) 76 polypeptide (an  
 CC alternatively spliced form of DNA polymerase kappa), appearing as  
 CC ABO07327 (encoded by the cDNA appearing as ACD13492). The gene for  
 CC POLKAPPA is located on chromosome 5q12-13 in a head-head arrangement with  
 CC the gene encoding Goodpasture antigen binding protein (GPBP or collagen 4  
 CC alpha 3 binding protein (COL4A3BP)), associated with autoimmune diseases  
 CC such as Goodpasture's disease and cutaneous lupus) i.e. has a bi-  
 CC directional promoter. Also included are a recombinant expression vector  
 CC comprising the polk76 cDNA, a host cell transfected with the vector,  
 CC detecting (M1) polk76 (comprising providing a protein sample to be  
 CC screened, contacting the protein sample to be screened with an anti-  
 CC polk76 antibody and detecting the formation of an antibody-polypeptide  
 CC complexes, where the presence of the antibody-polypeptide complexes  
 CC indicates the presence of polk76), detecting (M2) the polk76 nucleic acid  
 CC in a sample (comprising contacting the sample with one or more polk76 PCR  
 CC primer, carrying out PCR to generate PCR products, and identifying the  
 CC polk76-specific PCR), detecting an autoimmune condition in a patient  
 CC (comprising providing a tissue or body fluid sample from the patient,  
 CC providing a control tissue or body fluid sample in which no autoimmune  
 CC condition is present, and detecting an increase in pol k76 RNA expression  
 CC in the tissue of body fluid samples compared to the control sample, where  
 CC the increase indicates the presence of an autoimmune condition) and  
 CC treating (M3) a patient with an autoimmune disorder or cancer by  
 CC modifying the expression or activity of pol k76 in the patient. Modifying  
 CC the expression or activity of polk76 or polk76 nucleic acid, such as by  
 CC increasing or decreasing their expression or activity using antibodies or  
 CC antisense therapy, is useful for treating an autoimmune disorder or  
 CC cancer. The present sequence is a PCR fragment representing part (or all)



CC of the intergenic region or bi-directional promoter of the polkappa/GPBP  
CC genes  
XX  
SQ Sequence 771 BP; 123 A; 253 C; 209 G; 186 T; 0 U; 0 Other;  
Query Match 100.0%; Score 140; DB 9; Length 771;  
Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTTCGCGTGTCTCGTCAGACGCCGGGGGAGCG 60  
DB 195 GGGTTCCGGAGGAGGATCCCGAAGGCTTCGCGTGTCTCGTCAGACGCCGGGGGAGCG 136  
QY 61 GGGCGGGGAGTGTAGTGGGGAGATGGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 120  
DB 135 GGGCGGGGAGTGTAGTGGGGAGATGGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 76  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
DB 75 GAGGGTAAATAGTGGGCCAG 56  
RESULT 12  
ACD13466  
ID ACD13466 standard; DNA; 771 BP.  
XX  
AC ACD13466;  
DT 14-AUG-2003 (first entry)  
XX  
XX Human GPBP/COL4A3BP 771bp PCR fragment.  
XX  
KW Human; ds; Goodpasture antigen binding protein; GPBP; COL4A3BP;  
KW collagen 4 alpha 3 binding protein; DNA polymerase kappa; Pol kappa;  
KW Goodpasture disease; cutaneous lupus; polk76; bi-directional promoter;  
KW autoimmune disease; cancer; antisense therapy; chromosome 5q12-13.  
XX  
OS Homo sapiens.  
XX  
XX US2003027165-A1.  
XX  
XX 06-FEB-2003.  
XX  
XX 07-DEC-2001; 2001US-00010920.  
XX  
XX 08-DEC-2000; 2000US-0254649P.  
XX  
XX (SAUS/) SAUS J.  
XX  
XX Saus J;  
XX  
XX WPI; 2003-479531/45.  
XX  
XX New isolated DNA polymerase, pol kappa 76, useful in identifying  
XX autoimmune disorders and in treating cancer and autoimmune disorders by  
XX modifying its expression.  
XX  
XX Example; Page 20-21; 54pp; English.  
XX  
XX The invention relates to an isolated pol kappa (k) 76 polypeptide (an  
XX alternatively spliced form of DNA polymerase kappa), appearing as  
XX ABO07327 (encoded by the cDNA appearing as ACD13492). The gene for  
XX POLKAPPA is located on chromosome 5q12-13 in a head-head arrangement with  
XX the gene encoding Goodpasture antigen binding protein (GPBP or collagen 4  
XX alpha 3 binding protein (COL4A3bp), associated with autoimmune diseases  
XX such as Goodpasture's disease and cutaneous lupus) i.e. has a bi-  
XX directional promoter. Also included are a recombinant expression vector  
XX comprising the polk76 cDNA, a host cell transfected with the vector,  
XX detecting (M1) polk76 (comprising providing a protein sample to be  
XX screened, contacting the protein sample to be screened with an anti-  
XX polk76 antibody and detecting the formation of an antibody-polypeptide  
XX complexes, where the presence of the antibody-polypeptide complexes  
XX indicates the presence of polk76), detecting (M2) the polk76 nucleic acid

CC in a sample (comprising contacting the sample with one or more polk76 PCR  
CC primer, carrying out PCR to generate PCR products, and identifying the  
CC polk76-specific PCR), detecting an autoimmune condition in a patient  
CC (comprising providing a tissue or body fluid sample from the patient,  
CC providing a control tissue or body fluid sample in which no autoimmune  
CC condition is present, and detecting an increase in pol k76 RNA expression  
CC in the tissue of body fluid samples compared to the control sample, where  
CC the increase indicates the presence of an autoimmune condition) and  
CC treating (M3) a patient with an autoimmune disorder or cancer by  
CC modifying the expression or activity of pol k76 in the patient. Modifying  
CC the expression or activity of polk76 or polk76 nucleic acid, such as by  
CC increasing or decreasing their expression or activity using antibodies or  
CC antisense therapy, is useful for treating an autoimmune disorder or  
CC cancer. The present sequence is a PCR fragment representing part (or all)  
CC of the intergenic region or bi-directional promoter of the polkappa/GPBP  
XX genes  
SQ Sequence 771 BP; 186 A; 209 C; 253 G; 123 T; 0 U; 0 Other;  
Query Match 100.0%; Score 140; DB 9; Length 771;  
Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTTCGCGTGTCTCGTCAGACGCCGGGGGAGCG 60  
DB 577 GGGTTCCGGAGGAGGATCCCGAAGGCTTCGCGTGTCTCGTCAGACGCCGGGGGAGCG 636  
QY 61 GGGCGGGGAGTGTAGTGGGGAGATGGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 120  
DB 637 GGGCGGGGAGTGTAGTGGGGAGATGGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 696  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
DB 697 GAGGGTAAATAGTGGGCCAG 716  
RESULT 13  
ADA97869/c  
ID ADA97869 standard; DNA; 771 BP.  
XX  
XX ADA97869;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Human tumour necrosis factor (TNF) inducible promoter #3.  
XX  
XX Human; tumour necrosis factor inducible promoter; TNF;  
XX autoimmune disorder; cancer; ds; promoter; immunosuppressive; cytostatic.  
XX  
XX Homo sapiens.  
XX  
XX US2003082745-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 07-DEC-2001; 2001US-00008721.  
XX  
XX 08-DEC-2000; 2000US-0254649P.  
XX  
XX (SAUS/) SAUS J.  
XX  
XX Saus J;  
XX  
XX WPI; 2003-606062/57.  
XX  
XX New tumor necrosis factor inducible promoters, useful for identifying  
XX promoters that are regulated by tumor necrosis factor, or for identifying  
XX candidate compounds for treating or preventing autoimmune disorders or  
XX cancer.  
XX  
XX Claim 1; Fig 1; 57pp; English.  
XX  
XX The invention relates to a tumour necrosis factor (TNF) inducible

CC promoter. Also disclosed are an expression vector comprising one or more  
CC tumour necrosis factor inducible promoters and a recombinant host cell  
CC transfected with one or more expression vectors. The TNF inducible  
CC promoters, expression vectors and host cells are useful for identifying  
CC promoters that are regulated by tumour necrosis factor or for identifying  
CC candidate compounds for treating or preventing autoimmune disorders or  
CC cancer. This sequence represents a tumour necrosis factor inducible  
CC promoter of the invention.

XX SQ Sequence 771 BP; 123 A; 253 C; 209 G; 186 T; 0 U; 0 Other;  
Query Match 100.0%; Score 140; DB 9; Length 771;  
Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGAGGAGGATCCGAAGGCTCGCGTGTCTCGCTCAGACGCCGGAGGGGACG 60  
DB 195 GGGTTCGGAGGAGGATCCGAAGGCTCGCGTGTCTCGCTCAGACGCCGGAGGGGACG 136

QY 61 GGGCGGGGAGTAGTGGGGGAGATGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 120  
DB 135 GGGCGGGGAGTAGTGGGGGAGATGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 76

QY 121 GAGGGTAAATAGTGGGCCAG 140  
DB 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 14  
ADA97868  
ID ADA97868 standard; DNA; 771 BP.  
AC ADA97868;  
XX  
XX 20-NOV-2003 (first entry)  
XX Human tumour necrosis factor (TNF) inducible promoter #2.  
XX Human; tumour necrosis factor inducible promoter; TNF;  
KW autoimmune disorder; cancer; ds; promoter; immunosuppressive; cytostatic.  
XX OS. Homo sapiens.  
XX US2003082745-A1.  
XX PD 01-MAY-2003.  
XX 07-DEC-2001; 2001US-00008721.  
XX 08-DEC-2000; 2000US-0254649P.  
XX (SAUS/) SAUS J.  
XX Saus J;  
XX WPI; 2003-606062/57.  
XX New tumor necrosis factor inducible promoters, useful for identifying  
PT promoters that are regulated by tumor necrosis factor, or for identifying  
PT candidate compounds for treating or preventing autoimmune disorders or  
PT cancer.  
XX Claim 1; Fig 1; 57pp; English.  
XX The invention relates to a tumour necrosis factor (TNF) inducible  
CC promoter. Also disclosed are an expression vector comprising one or more  
CC tumour necrosis factor inducible promoters and a recombinant host cell  
CC transfected with one or more expression vectors. The TNF inducible  
CC promoters, expression vectors and host cells are useful for identifying  
CC promoters that are regulated by tumour necrosis factor or for identifying  
CC candidate compounds for treating or preventing autoimmune disorders or  
CC cancer. This sequence represents a tumour necrosis factor inducible  
CC promoter of the invention.

XX SQ Sequence 771 BP; 186 A; 209 C; 253 G; 123 T; 0 U; 0 Other;  
Query Match 100.0%; Score 140; DB 9; Length 771;  
Best Local Similarity 100.0%; Pred. No. 6.7e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGAGGAGGATCCGAAGGCTCGCGTGTCTCGCTCAGACGCCGGAGGGGACG 60  
DB 577 GGGTTCGGAGGAGGATCCGAAGGCTCGCGTGTCTCGCTCAGACGCCGGAGGGGACG 636

QY 61 GGGCGGGGAGTAGTGGGGGAGATGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 120  
DB 637 GGGCGGGGAGTAGTGGGGGAGATGGAGACGAAGGGGAGGGAAGGACAGGGGAGGG 696

QY 121 GAGGGTAAATAGTGGGCCAG 140  
DB 697 GAGGGTAAATAGTGGGCCAG 716

RESULT 15  
AAD41041  
ID AAD41041 standard; DNA; 955 BP.  
XX  
XX AAD41041;  
XX 30-OCT-2002 (first entry)  
XX Human 955-bp PCR product DNA.  
XX Human; tumour necrosis factor; TNF; promoter; autoimmune disorder;  
KW cancer; therapy; ds.  
XX OS. Homo sapiens.  
XX Key  
FH exon  
FT 1. .443  
FT /\*tag= a  
FT /note= "Exon sequence of POLK contained in HeLa 4.1"  
FT complement(1. .21)  
FT /\*tag= b  
FT /bound\_moiety= "Primer ON-GPBP-18m"  
FT complement(443)  
FT /\*tag= c  
FT /note= "Transcriptional start site"  
FT 567. .735  
FT /\*tag= d  
FT /note= "Exon 1 of POLK"  
FT 726. .865  
FT /\*tag= e  
FT /note= "140bp present in Sprompolk and SprompGPBP"  
FT 851. .856  
FT /\*tag= f  
FT 858. .955  
FT /\*tag= g  
FT /note= "Part of exon 1 of COL4A3BP"  
FT 858  
FT /\*tag= h  
FT /note= "Transcriptional start site"  
FT 930. .955  
FT /\*tag= i  
FT /bound\_moiety= "ON-GPBP-6c"  
XX WO200246433-A2.  
XX 13-JUN-2002.  
XX 07-DEC-2001; 2001WO-EP014412.  
XX 08-DEC-2000; 2000US-0254649P.  
XX (SAUS/) SAUS J.  
XX

PI Saus J;  
XX WPI; 2002-519670/55.  
XX  
XX Novel tumor necrosis-factor inducible promoter useful for identifying  
PT candidate compounds for treating/preventing autoimmune disorders/cancer,  
PT or for identifying promoters that are regulated by tumor necrosis factor.  
XX  
XX Claim 1; Page 66-67; 95pp; English.  
XX  
XX The invention relates to a tumour necrosis-factor TNF inducible promoter.  
CC The invention is useful for identifying candidate TNF inducible promoters  
CC by aligning a test sequence consisting of a nucleic acid sequence with a  
CC comparison sequence selected from the invention, using a gap opening  
CC penalty of 50 and a gap extension penalty of 3 to define a test  
CC alignment, shuffling the nucleic sequence of the test sequence at least  
CC one hundred times, while maintaining its length and composition, to  
CC produce a series of randomised sequences, aligning the randomised  
CC sequences with the comparison sequence using a gap opening penalty of 50  
CC and a gap extension penalty of 3, to produce a series of randomised  
CC alignments, determining an average alignment quality of the randomised  
CC alignments, where the average alignment quality of the randomised  
CC alignments represent an alignment expected by chance, comparing the test  
CC alignment with the average alignment quality of the randomised alignments  
CC and identifying a test alignment with a probability value of less than  
CC 0.05 that the alignment is obtained by chance as a candidate TNF  
CC inducible promoter. The invention is useful for identifying candidate  
CC compounds for treating or preventing autoimmune disorders or cancer. The  
CC present sequence is human 955-bp PCR product DNA  
XX  
SQ Sequence 955 BP; 236 A; 257 C; 271 G; 191 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 140; DB 6; Length 955;  
Best Local Similarity 100.0%; Pred. No. 6.8e-24;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GGGTTCCGGAGGAGGATCCCGAGGCTCGCGTGTCTCGCTCAGACGCCGGGAGGGGACG 60  
Db |  
726 GGGTTCCGGAGGAGGATCCCGAGGCTCGCGTGTCTCGCTCAGACGCCGGGAGGGGACG 785  
|  
  
Qy 61 GGGCGGGAGTGTGGGGGAGATGGGAGGACGAAGGGGGGAAAGGACAGGGGAGGG 120  
Db |  
786 GGGCGGGAGTGTGGGGGAGATGGGAGGACGAAGGGGGGAAAGGACAGGGGAGGG 845  
|  
  
Qy 121 GAGGGTAAATAGTGGGCAG 140  
Db |  
846 GAGGGTAAATAGTGGGCAG 865  
|

Search completed: June 12, 2005, 20:48:38  
Job time : 282 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2005, 20:07:37 ; Search time 106 Seconds  
(without alignments)  
2161.120 Million cell updates/sec

Title: US-10-008-721-6  
Perfect score: 140  
Sequence: 1 ggggttcggaggagatcc.....gagggttaaatagtggtccag 140

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	49	35.0	7218	1	US-08-232-463-14
C 2	44.6	31.9	101951	4	US-09-949-016-15648
C 3	44.4	31.7	46823	4	US-09-949-016-12723
C 4	44.4	31.7	46940	4	US-09-949-016-16252
C 5	44.4	31.7	99748	4	US-09-949-016-11990
C 6	44.4	31.7	99749	4	US-09-949-016-16518
C 7	44	31.4	500	4	US-09-866-108A-15742
C 8	43.8	31.3	601	4	US-09-949-016-135107
C 9	43.8	31.3	117807	4	US-09-949-016-15525
C 10	43.2	30.9	865	4	US-09-270-767-11042
C 11	43.2	30.9	16782	4	US-09-949-016-17291
C 12	43	30.7	320	3	US-09-165-264-14
C 13	42.8	30.6	320	3	US-09-165-264-11
C 14	42.4	30.3	320	3	US-09-165-264-13
C 15	42.4	30.3	53526	3	US-08-658-136-2
C 16	42.4	30.3	53577	3	US-08-658-136-1
C 17	42.4	30.3	119032	4	US-09-949-016-12160
C 18	42.4	30.3	119032	4	US-09-949-016-17268
C 19	42.2	30.1	320	3	US-09-165-264-7
C 20	41.8	29.9	1469	4	US-09-865-879-12
C 21	41.6	29.7	318	3	US-09-165-264-12
C 22	41.2	29.4	319	3	US-09-165-264-8
C 23	41.2	29.4	264665	4	US-09-949-016-13747
C 24	41	29.3	253345	4	US-09-949-016-12656
C 25	41	29.3	233364	4	US-09-949-016-13639
C 26	40.6	29.0	936	4	US-09-270-767-4464
C 27	40.6	29.0	936	4	US-09-270-767-19746

c 28	40.6	29.0	9293	4	US-09-949-016-16801
c 29	40.6	29.0	37875	4	US-09-949-016-13182
c 30	40.4	28.9	204	4	US-09-107-433-2184
c 31	40.4	28.9	209	4	US-09-107-433-185
c 32	40.4	28.9	282	4	US-09-107-433-184
c 33	40.4	28.9	308	4	US-09-107-433-1723
c 34	40.2	28.7	83617	4	US-09-107-433-1970
c 35	40.2	28.7	83617	4	US-09-949-016-12254
c 36	40.2	28.7	87870	4	US-09-949-016-14461
c 37	39.8	28.4	390890	4	US-09-949-016-14720
c 38	39.6	28.3	12001	1	US-08-458-568A-11
c 39	39.6	28.3	60950	4	US-09-949-016-14080
c 40	39.6	28.3	256287	4	US-09-949-016-14608
c 41	39.4	28.1	601	4	US-09-949-016-127819
c 42	39.4	28.1	601	4	US-09-949-016-127820
c 43	39.4	28.1	601	4	US-09-949-016-128156
c 44	39.4	28.1	601	4	US-09-949-016-128157
c 45	39.4	28.1	304533	4	US-09-949-016-15371

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMWU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
; US-08-232-463-14

Query Match 35.0% Score 49; DB 1; Length 7218;

```

Best Local Similarity 3.0%; Pred. No. 0.00063;
Matches 4; Conservative 103; Mismatches 28; Indels 0; Gaps 0;

QY 2 GGTTCGGAGGAGATCCGGAAGCTCGCGTTCAGACGCCGGGAGGGGACGG 61
Db 1439 GGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1380
QY 62 GCGCGGGAGTAGTGGGGAGAAATGGGAGGACGAAGGGAGGGAAGGACAGGGAGGG 121
Db 1379 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1320
QY 122 AGCGTAAATAGTCGG 136
Db 1319 RRRRRRRRRRRR 1305

RESULT 2
US-09-949-016-15648/c
; Sequence 15648, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15648
; LENGTH: 101951
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15648

Query Match 31.9%; Score 44.6; DB 4; Length 101951;
Best Local Similarity 71.1%; Pred. No. 0.013;
Matches 59; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 43 GACCGCGGAGGGGACGGCGCGGAGTAGTCGGGGAGAGATCGGAGGACGAGGGAGG 102
Db 28062 GAGGGAGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGG 28003
QY 103 GGAAAGGACAGGGGAGGGAGGG 125
Db 28002 GGGAGGGGAGGGGAGGGAGCG 27980

RESULT 3
US-09-949-016-12723
; Sequence 12723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 12723
; LENGTH: 46823
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(46823)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12723

Query Match 31.7%; Score 44.4; DB 4; Length 46823;
Best Local Similarity 58.2%; Pred. No. 0.013;
Matches 78; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 7 GGGAGGAGGATCCGGAAGCTCGCGTTCAGACGCCGGGAGGGGACGGGGCGG 66
Db 16710 GGGAGGAGGAGGACAAAGAGGAAGAGGAGGAGGAGGATGGGGAGGAGGA 16769
QY 67 GGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGGAAAGGACAGGGGAGGGGT 126
Db 16770 GGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16829
QY 127 AAATAGTGGGCCAG 140
Db 16830 AAAAGGAGGAGGAG 16843

RESULT 4
US-09-949-016-16252
; Sequence 16252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16252
; LENGTH: 46940
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(46940)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16252

Query Match 31.7%; Score 44.4; DB 4; Length 46940;
Best Local Similarity 58.2%; Pred. No. 0.013;
Matches 78; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 7 GGGAGGAGGATCCGGAAGCTCGCGTTCAGACGCCGGGAGGGGACGGGGCGG 66
Db- 16710 GGGAGGAGGAGGACAAAGAGGAAGAGGAGGAGGAGGATGGGGAGGAGGA 16769
QY 67 GGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGGAAAGGACAGGGGAGGGGT 126
Db 16770 GGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16829
QY 127 AAATAGTGGGCCAG 140
Db 16830 AAAAGGAGGAGGAG 16843

```

RESULT 5  
US-09-949-016-11990  
; Sequence 11990, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11990  
; LENGTH: 99748  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11990

Query Match 31.7%; Score 44.4; DB 4; Length 99748;  
Best Local Similarity 61.0%; Pred. No. 0.015;  
Matches 72; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 7 GGGAGGAGGATCCGAGAGGTCGCGTGTCTCAGACGCCGGGAGGGGAGCGGGCGG 66  
Db 70064 GAGGAGGATAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 70123

QY 67 GGGAGTGTGGGAGAGTGGGAGGAGCGAAGCGGAGGAGGAGGAGGAGGAGGAGG 124  
Db 70124 AGAGCAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 70181

RESULT 6  
US-09-949-016-16518  
; Sequence 16518, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16518  
; LENGTH: 99749  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16518

Query Match 31.7%; Score 44.4; DB 4; Length 99749;  
Best Local Similarity 61.0%; Pred. No. 0.015;  
Matches 72; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 7 GGGAGGAGGATCCGAGAGGTCGCGTGTCTCAGACGCCGGGAGGGGAGCGGGCGG 66  
Db 70064 GAGGAGGATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 70123

QY 67 GGGAGTGTGGGAGAGTGGGAGGAGCGAAGCGGAGGAGGAGGAGGAGGAGGAGG 124

Db 70124 AGAGCAAGGAGGAGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 70181

RESULT 7  
US-09-866-108A-15742/c  
; Sequence 15742, Application US/09866108A  
; Patent No. 6686188  
; GENERAL INFORMATION:  
; APPLICANT: GU, Yizhong  
; APPLICANT: JI, Yonggang  
; APPLICANT: PENN, Sharon G.  
; APPLICANT: HANZEL, David K.  
; APPLICANT: RANK, David R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: SHANNON, Mark  
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE  
; FILE REFERENCE: AEOMICA-7  
; CURRENT APPLICATION NUMBER: US/09/866,108A  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 15755  
; SOFTWARE: Aescima Sequence Listing Engine  
; Patent No. 6686188  
; SEQ ID NO 15742  
; LENGTH: 500  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-866-108A-15742

Query Match 31.4%; Score 44; DB 4; Length 500;  
Best Local Similarity 59.7%; Pred. No. 0.009;  
Matches 74; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGGATCCGAGAGGTCGCGTGTCTCAGACGCCGGGAGGGGACG 60  
Db 137 GGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 78

QY 61 GGGCGGGGAGTGTAGTGGGGAGAGTGGAGGAGCAAGAGGGGAGGGGAGGAGGAGG 120  
Db 77 AGGGGAGGGAGGAGTGGAGGGGAGGGGAGGGGAGGAGGAGGAGGAGGAGGAGG 18

QY 121 GAGG 124  
Db 17 GAGG 14

RESULT 8  
US-09-949-016-135107/c  
; Sequence 135107, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 135107  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-135107

Query Match 31.3%; Score 43.8; DB 4; Length 601;  
Best Local Similarity 60.5%; Pred. No. 0.01;  
Matches 72; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 7 GGGAGGAGGATCCGAAAGCTCGCGTGTGCGTCAGACCCCGGAGGGGAGACGGGGCGG 66  
Db 470 GAGAGGGAGAGGAGAGAGGATAGCGCAGAGGGAGAGGATCGGAGAGGAGAGGAGAGG 411

QY 67 GGAGTAGTGTGGGGAGAGATGGAGACGACGAAGCGGGAGGAGGAGACAGCGGGAGGGAGGG 125  
Db 410 GAAAGGACAGGAGGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGG 352

RESULT 9  
US-09-949-016-15525/c  
; Sequence 15525, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 15525  
; LENGTH: 117807  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15525

Query Match 31.3%; Score 43.8; DB 4; Length 117807;  
Best Local Similarity 60.5%; Pred. No. 0.021;  
Matches 72; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 7 GGGAGGAGGATCCGAAAGCTCGCGTGTGCGTCAGACCCCGGAGGGGAGACGGGGCGG 66  
Db 54166 GAGAGGGAGAGGGAGAGGATAGCGCAGAGGGAGAGGATGGGAGAGGAGGAGGAGAGG 54107

QY 67 GGAGTAGTGGGGAGAGATGGAGAGACGAGCGGAGGGGAGAAAGACAGGGAGGGAGGG 125  
Db 54106 GAAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 54048

RESULT 10  
US-09-270-767-11042/c  
; Sequence 11042, Application US/09270767  
; Patent No. 6703491

; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11042  
; LENGTH: 865  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: n means any nucleotide  
US-09-270-767-11042

Query Match 30.9%; Score 43.2; DB 4; Length 865;  
Best Local Similarity 57.4%; Pred. No. 0.016;  
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTCAGACCCCGGAGGGGACG 60  
Db 276 GGG 217

QY 61 GGGCGGGGGAGTAGTGTGGGGAGAGATGGAGACCGAAGGGGAGGGGAGGAGGAGGAGGG 120  
Db 216 GGG 157

QY 121 GAGGGTAAATAGTGGG 136  
Db 156 GGGGGGGGGTGTGGG 141

RESULT 11  
US-09-949-016-17291  
; Sequence 17291, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 17291  
; LENGTH: 16782  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17291

Query Match 30.9%; Score 43.2; DB 4; Length 16782;  
Best Local Similarity 57.4%; Pred. No. 0.023;  
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTGCGTCAGACCCCGGAGGGGACG 60  
Db 1220 GGGGAGGGTAGGGGGGTGGGGGGGTAGGGGGGTGGGGAGGGGTAGGGGGGGTGGGGGGGTG 1279

QY 61 GGGCGGGGAGTAGTGTGGGGAGAGATGGAGACCGAAGGGGAGGGGAGGAGGAGGAGGGG 120  
Db 1280 GGGAGGGTAGGGGGTGGGGGGGTGGGGAGAGTAGGGGGGTGGGGAGAGTAGGGGGGTGGG 1339

QY 121 GAGGGTAAATAGTGGG 136  
Db 1340 GAGAGTAGGGGGTGGG 1355



Db 297 GGGGCGAGATTTCG 310

RESULT 14  
US-09-165-264-13  
; Sequence 13, Application US/09165264  
; Patent No. 6197510  
; GENERAL INFORMATION:  
; APPLICANT: Vinayagamoorthy, Thuraiayah  
; TITLE OF INVENTION: Multi-Loci Genomic Analysis  
; FILE REFERENCE: 44747  
; CURRENT APPLICATION NUMBER: US/09/165,264  
; CURRENT FILING DATE: 1998-10-01  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence  
US-09-165-264-13

Query Match 30.3%; Score 42.4; DB 3; Length 320;  
Best Local Similarity 56.4%; Pred.No.0.022;  
Matches 79; Conservative 0; Mismatches 61; Indels 0; Gaps 0

QY 1 GGGCTTCGGGAGGAGGATCCCGAAGGCTCGGCCTGTCCGTCAGACGCCGGGAGGGGGGACG 60  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 172 GG 231  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 61 GGGCGGGGACTAGTGGGGGAGAATGGAGAGCAGAAAGGGGAGGGGAAAAGACAAGGGGAGGG 120  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 232 GG 291  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 121 GAGGTAAATAGTGGGCCAG 140  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 292 GGGGGGGGGGACGTGGACGAG 311

RESULT 15  
US-08-658-136-2/c  
; Sequence 2, Application US/08658136  
; Patent No. 6071717  
; GENERAL INFORMATION:  
; APPLICANT: KLINGER, KATHERINE W  
; APPLICANT: LANDES, GREGORY M  
; APPLICANT: BURN, TIMOTHY C  
; APPLICANT: CONNORS, TIMOTHY D  
; APPLICANT: DACKOWSKI, WILLIAM  
; APPLICANT: GERMINO, GREGORY  
; APPLICANT: QIAN, FENG  
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: ONE MOUNTAIN ROAD  
; CITY: FRAMINGHAM  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/658,136  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LASSEN, ELIZABETH  
; REGISTRATION NUMBER: 31,845

```

; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-658-136-2

```

```

Query Match      30.3%; Score 42.4; DB 3; Length 53526;
Best Local Similarity 57.6%; Pred. No. 0.045;
Matches 76; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCCGGCGTGTCCGCTCAGACGCCGGGAGGGGACG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34341 GCGAAGGCTGGGGGGAAGAGGAAAGGCTAGGGGAGGGAGGGAGGGGCTAG 34282

QY 61 GGGCGGGGAGTAGTGGGGGAGAGTGGAGAGCAGAGAGGGGAGGGGAAAGGACAGGGGAGGG 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34281 GCGAGGGGAGGAGGGGAGGGGCTAGGGGAGGGAAGGGGAGGGGAGGGGAGGGGAGAGTG 34222

QY 121 GAGGTAATAG 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34221 GAGGGCACAGAG 34210

```

Search completed: June 12, 2005, 21:19:35  
Job time : 109 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2005, 20:20:48 ; Search time 329 Seconds  
(without alignments)

2637.898 Million cell updates/sec

Title: US-10-008-721-6

Perfect score: 140

Sequence: 1 ggggtcggaggaggatccc.....gagggtaaatagtggccag 140

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	140	14	US-10-010-920-6
2	140	100.0	140	14	US-10-010-920-7
3	140	100.0	140	14	US-10-008-721-6
4	140	100.0	140	14	US-10-008-721-7
5	140	100.0	771	14	US-10-010-920-4
6	140	100.0	771	14	US-10-010-920-5
7	140	100.0	771	14	US-10-008-721-4

c	8	140	100.0	771	14	US-10-008-721-5	Sequence 5, Appli
	9	140	100.0	955	14	US-10-010-920-3	Sequence 3, Appli
	10	140	100.0	955	14	US-10-008-721-3	Sequence 3, Appli
c	11	50.4	36.0	424	19	US-10-437-963-49515	Sequence 49515, A
	12	48	34.3	985	20	US-10-425-115-91623	Sequence 91623, A
	13	47.6	34.0	408	19	US-10-021-323-5007	Sequence 5007, Ap
	14	47	33.6	588	19	US-10-021-323-9377	Sequence 9377, Ap
c	15	46.6	33.3	406	19	US-10-021-323-14371	Sequence 14371, A
c	16	46.6	33.3	891	20	US-10-425-115-146045	Sequence 146045,
	17	46.6	33.3	20345	17	US-10-292-798-705	Sequence 705, App
c	18	46.4	33.1	675	20	US-10-425-115-45871	Sequence 45871, A
c	19	46.2	33.0	92219	19	US-10-322-281-805	Sequence 805, App
c	20	46	32.9	632	20	US-10-425-115-16028	Sequence 16028, A
	21	45.8	32.7	804	18	US-10-424-599-100535	Sequence 100535,
c	22	45.6	32.6	614	20	US-10-425-115-94057	Sequence 94057, A
c	23	45.6	32.6	629	19	US-10-021-323-9375	Sequence 9375, Ap
c	24	45.2	32.3	673	20	US-10-425-115-11145	Sequence 11145, A
c	25	45.2	32.3	1065	9	US-09-804-682-33	Sequence 33, Appl
	26	45	32.1	263	19	US-10-021-323-16665	Sequence 16665, A
c	27	45	32.1	668	19	US-10-437-963-3757	Sequence 3757, Ap
c	28	45	32.1	766	20	US-10-425-115-82176	Sequence 82176, A
c	29	45	32.1	925	19	US-10-437-963-44536	Sequence 44536, A
	30	45	32.1	1211	20	US-10-425-115-93385	Sequence 93385, A
	31	45	32.1	1385	17	US-10-398-221-1379	Sequence 1379, Ap
	32	45	32.1	1385	17	US-10-398-221-2930	Sequence 2930, Ap
c	33	44.8	32.0	294	20	US-10-425-115-79029	Sequence 79029, A
	34	44.8	32.0	447	19	US-10-021-323-10481	Sequence 10481, A
c	35	44.8	32.0	492	18	US-10-424-599-59619	Sequence 59619, A
c	36	44.8	32.0	545	19	US-10-437-963-56258	Sequence 56258, A
c	37	44.8	32.0	580	18	US-10-424-599-2320	Sequence 2320, Ap
	38	44.8	32.0	636	20	US-10-425-115-168126	Sequence 168126,
c	39	44.8	32.0	1087	18	US-10-424-599-81404	Sequence 81404, A
	40	44.8	32.0	189817	19	US-10-741-601-5660	Sequence 5660, Ap
	41	44.8	32.0	189817	21	US-10-741-600-17685	Sequence 17685, A
	42	44.6	31.9	522	13	US-10-027-632-134954	Sequence 134954,
	43	44.6	31.9	522	13	US-10-027-632-134955	Sequence 134955,
	44	44.6	31.9	522	17	US-10-027-632-134954	Sequence 134954,
	45	44.6	31.9	522	17	US-10-027-632-134955	Sequence 134955,

ALIGNMENTS

RESULT 1  
US-10-010-920-6  
; Sequence 6, Application US/10010920  
; Publication NO. US20030027165A1  
; GENERAL INFORMATION:  
; APPLICANT: Saus, Juan  
; TITLE OF INVENTION: Alternatively spliced polk nucleotide and amino acid sequences  
; TITLE OF INVENTION: and methods for using  
; FILE REFERENCE: 98,723-E3  
; CURRENT APPLICATION NUMBER: US/10/010,920  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,649  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 140  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-010-920-6

Query Match	100.0%	Score	140;	DB	14;	Length	140;
Best Local Similarity	100.0%	Pred. NO.	3.5e-29;				
Matches	140;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	GGGTTTCGGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGGGGGGACG	60				
Db	1	GGGTTTCGGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGGGGGGACG	60				
Qy	61	GGGCGGGGAGTAGTGTGGGGGAGATGGGAGGACGAAAGGGGGGAGGACAGGGGGGGG	120				

Db 61 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 120  
|  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
|  
Db 121 GAGGGTAAATAGTGGGCCAG 140  
|

## RESULT 2

US-10-010-920-7/c  
; Sequence 7, Application US/10010920  
; Publication No. US20030027165A1  
; GENERAL INFORMATION:  
; APPLICANT: Saus, Juan  
; TITLE OF INVENTION: Alternatively spliced polk nucleotide and amino acid sequences  
; TITLE OF INVENTION: and methods for using  
; FILE REFERENCE: 98,723-E3  
; CURRENT APPLICATION NUMBER: US/10/010,920  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,649  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 140  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-010-920-7

Query Match 100.0%; Score 140; DB 14; Length 140;  
Best Local Similarity 100.0%; Pred. No. 3.5e-29;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60  
|  
Db 140 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 81  
|  
QY 61 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 120  
|  
Db 80 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 21  
|  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
|  
Db 20 GAGGGTAAATAGTGGGCCAG 1  
|

## RESULT 3

US-10-008-721-6  
; Sequence 6, Application US/10008721  
; Publication No. US20030082745A1  
; GENERAL INFORMATION:  
; APPLICANT: Saus, Juan  
; TITLE OF INVENTION: TNF-Inducible Promoters and Methods for Using  
; FILE REFERENCE: 98,723-E1  
; CURRENT APPLICATION NUMBER: US/10/008,721  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,649  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 140  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-008-721-6

Query Match 100.0%; Score 140; DB 14; Length 140;  
Best Local Similarity 100.0%; Pred. No. 3.5e-29;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60  
|  
Db 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60  
|

QY 61 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 120  
|  
Db 61 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 120  
|  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
|  
Db 121 GAGGGTAAATAGTGGGCCAG 140  
|

## RESULT 4

US-10-008-721-7/c  
; Sequence 7, Application US/10008721  
; Publication No. US20030082745A1  
; GENERAL INFORMATION:  
; APPLICANT: Saus, Juan  
; TITLE OF INVENTION: TNF-Inducible Promoters and Methods for Using  
; FILE REFERENCE: 98,723-E1  
; CURRENT APPLICATION NUMBER: US/10/008,721  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,649  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 140  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-008-721-7

Query Match 100.0%; Score 140; DB 14; Length 140;  
Best Local Similarity 100.0%; Pred. No. 3.5e-29;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60  
|  
Db 140 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 81  
|  
QY 61 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 120  
|  
Db 80 GGGCGGGAGTACTGGGGAGAAATGGGAGACGAAGGGAGGGGAAAGACACGGGAGGG 21  
|  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
|  
Db 20 GAGGGTAAATAGTGGGCCAG 1  
|

## RESULT 5

US-10-010-920-4  
; Sequence 4, Application US/10010920  
; Publication No. US20030027165A1  
; GENERAL INFORMATION:  
; APPLICANT: Saus, Juan  
; TITLE OF INVENTION: Alternatively spliced polk nucleotide and amino acid sequences  
; TITLE OF INVENTION: and methods for using  
; FILE REFERENCE: 98,723-E3  
; CURRENT APPLICATION NUMBER: US/10/010,920  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,649  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 771  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-010-920-4

Query Match 100.0%; Score 140; DB 14; Length 771;  
Best Local Similarity 100.0%; Pred. No. 2.6e-29;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGGATCCCGAAGGCTCGCGGTGTCGCGTCAGACGCCGGAGGGGACG 60  
|

Db 577 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 636  
 Qy 61 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 120  
 Db 637 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 696  
 Qy 121 GAGGGTAAATAGTGGGCCAG 140  
 Db 697 GAGGGTAAATAGTGGGCCAG 716

RESULT 6  
 US-10-010-920-5/c  
 ; Sequence 5, Application US/10010920  
 ; Publication No. US20030027165A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saus, Juan  
 ; TITLE OF INVENTION: Alternatively spliced polk nucleotide and amino acid sequences  
 ; FILE REFERENCE: 98,723-E3  
 ; CURRENT APPLICATION NUMBER: US/10/010,920  
 ; CURRENT FILING DATE: 2001-12-07  
 ; PRIOR APPLICATION NUMBER: 60/254,649  
 ; PRIOR FILING DATE: 2000-12-08  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 771  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-010-920-5

Query Match 100.0%; Score 140; DB 14; Length 771;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-29;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 60  
 Db 195 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 136  
 Qy 61 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 120  
 Db 135 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 76  
 Qy 121 GAGGGTAAATAGTGGGCCAG 140  
 Db 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 7  
 US-10-008-721-4  
 ; Sequence 4, Application US/10008721  
 ; Publication No. US20030082745A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saus, Juan  
 ; TITLE OF INVENTION: TNF-Inducible Promoters and Methods for Using  
 ; FILE REFERENCE: 98,723-E1  
 ; CURRENT APPLICATION NUMBER: US/10/008,721  
 ; CURRENT FILING DATE: 2001-12-07  
 ; PRIOR APPLICATION NUMBER: 60/254,649  
 ; PRIOR FILING DATE: 2000-12-08  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 4  
 ; LENGTH: 771  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-008-721-4

Query Match 100.0%; Score 140; DB 14; Length 771;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-29;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 60  
 Db 577 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 636  
 Qy 61 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 120  
 Db 637 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 696  
 Qy 121 GAGGGTAAATAGTGGGCCAG 140  
 Db 697 GAGGGTAAATAGTGGGCCAG 716

RESULT 8  
 US-10-008-721-5/c  
 ; Sequence 5, Application US/10008721  
 ; Publication No. US20030082745A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saus, Juan  
 ; TITLE OF INVENTION: TNF-Inducible Promoters and Methods for Using  
 ; FILE REFERENCE: 98,723-E1  
 ; CURRENT APPLICATION NUMBER: US/10/008,721  
 ; CURRENT FILING DATE: 2001-12-07  
 ; PRIOR APPLICATION NUMBER: 60/254,649  
 ; PRIOR FILING DATE: 2000-12-08  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 771  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-008-721-5

Query Match 100.0%; Score 140; DB 14; Length 771;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-29;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 60  
 Db 195 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGTGTCCGTCAGACGCCGGAGGGGACG 136  
 Qy 61 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 120  
 Db 135 GGGCGGGAGTGTAGTGGGGAGGAGTGGAGGACGAAGAGGGGGAAGGACAGGGGAGGG 76  
 Qy 121 GAGGGTAAATAGTGGGCCAG 140  
 Db 75 GAGGGTAAATAGTGGGCCAG 56

RESULT 9  
 US-10-010-920-3  
 ; Sequence 3, Application US/10010920  
 ; Publication No. US20030027165A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saus, Juan  
 ; TITLE OF INVENTION: Alternatively spliced polk nucleotide and amino acid sequences  
 ; FILE REFERENCE: 98,723-E3  
 ; CURRENT APPLICATION NUMBER: US/10/010,920  
 ; CURRENT FILING DATE: 2001-12-07  
 ; PRIOR APPLICATION NUMBER: 60/254,649  
 ; PRIOR FILING DATE: 2000-12-08  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 955  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-010-920-3

Query Match 100.0%; Score 140; DB 14; Length 955;

Best Local Similarity 100.0%; Pred. No. 2.5e-29; Mismatches 0; Indels 0; Gaps 0;  
Matches 140; Conservative 0;

QY 1 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 60  
Db |||||  
726 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 785  
QY 61 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 120  
Db |||||  
786 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 845  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
Db 846 GAGGGTAAATAGTGGGCCAG 865

RESULT 10  
US-10-008-721-3  
; Sequence 3, Application US/10008721  
; Publication No. US20030082745A1  
; GENERAL INFORMATION:  
; APPLICANT: Saus, Juan  
; TITLE OF INVENTION: TNF-Inducible Promoters and Methods for Using  
; FILE REFERENCE: 98,723-E1  
; CURRENT APPLICATION NUMBER: US/10/008,721  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,649  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 955  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-008-721-3

Query Match 100.0%; Score 140; DB 14; Length 955;  
Best Local Similarity 100.0%; Pred. No. 2.5e-29; Mismatches 0; Indels 0; Gaps 0;  
Matches 140; Conservative 0;

QY 1 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 60  
Db 726 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 785  
QY 61 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 120  
Db 786 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 845  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
Db 846 GAGGGTAAATAGTGGGCCAG 865

RESULT 11  
US-10-437-963-49515/c  
; Sequence 49515, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 49515

; LENGTH: 424  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_52092C.1  
US-10-437-963-49515

Query Match 36.0%; Score 50.4; DB 19; Length 424;  
Best Local Similarity 60.0%; Pred. No. 9.1e-05; Mismatches 56; Indels 0; Gaps 0;  
Matches 84; Conservative 0;

QY 1 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 60  
Db 270 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 211  
QY 61 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 120  
Db 210 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 151  
QY 121 GAGGGTAAATAGTGGGCCAG 140  
Db 150 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 131

RESULT 12  
US-10-425-115-91623  
; Sequence 91623, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 91623  
; LENGTH: 985  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(985)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_183557C.1  
US-10-425-115-91623

Query Match 34.3%; Score 48; DB 20; Length 985;  
Best Local Similarity 59.6%; Pred. No. 0.00035; Mismatches 81; Conservative 0; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 60  
Db 536 GGGTTCGGGAGGAGATCCCGAAGGCTCGCGCTGTTCGCTCAGACGCCGGAGGGGAGCG 595  
QY 61 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 120  
Db 596 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 655  
QY 121 GAGGGTAAATAGTGGG 136  
Db 656 GGGCGGGGAGTACTGGGGAGAGATCGGAGCAGCAAGGGGAGGGAAGGACAGGGGAGGG 671

RESULT 13  
US-10-021-323-5007  
; Sequence 5007, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:

```

; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021.323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 5007
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3826-001-Q1-N6-H3
US-10-021-323-5007

```

```

Query Match      34.0%; Score 47.6; DB 19; Length 408;
Best Local Similarity 69.1%; Pred. No. 0.00053;
Matches 65; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 43 GAGCCCGGAGGGGACGCGCGGGGAGTAGTGGGGGAGAGTGGGAGGACGAGAGGGAGG 102
Db 142 GAAGAGGGGGGGGCAAGAGGGGGGGGGGAGAGGGGGGGGGGGGGGGGGGGGGGG 201

Qy 103 GGAAGGACAGGGGAGGGGAGGGGTAATAGTGGG 136
Db 202 GGAAGGGGAAGGGAAGAGGGGACCTGTGGG 235

```

```

RESULT 14
US-10-021-323-9377
; Sequence 9377, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021.323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 9377
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(588)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3828-025-Q6-K6-D5
US-10-021-323-9377

```

```

Query Match      33.6%; Score 47; DB 19; Length 588;
Best Local Similarity 62.2%; Pred. No. 0.00073;
Matches 74; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 7 GGGAGGAGGATCCCGAGGGCTCGCGGTCTCAGACGCCGGGAGGGAGCGGGGGGG 66
Db 397 GGGAGGGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 456

Qy 67 GGAGTAGTGGGGGAGATGGGAGGACGAGGGGGGGGGAAGAGACAGGGAGGGGAGG 125
Db 457 GAGGGGGGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 515

```

```

RESULT 15
US-10-021-323-14271/c
; Sequence 14271, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021.323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14271
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(406)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3829-031-Q6-N6-H2
US-10-021-323-14271

```

```

Query Match      33.3%; Score 46.6; DB 19; Length 406;
Best Local Similarity 60.8%; Pred. No. 0.001;
Matches 76; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 GGGTTCCGGAGGAGGATCCCGAAGGCTCGCGGTCTCGCGTCAGACGCCGGGAGGGGACG 60
Db 321 GGGGGGGGGGGAGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 262

Qy 61 GGGCGGGGAGTAGTGGGGGAGAGTGGGAGGACGAGAGGGGGGAGGAGGAGGAGGAGG 120
Db 261 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 202

Qy 121 GAGGG 125
Db 201 GGGGG 197

```

Search completed: June 12, 2005, 21:25:09  
Job time : 331 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2005, 19:55:18 ; Search time 1736 Seconds  
(without alignments)  
3069.699 Million cell updates/sec

Title: US-10-008-721-6  
Perfect score: 140  
Sequence: 1 ggggtcggaggaggatccc.....gagggtaaatagtgggcccag 140

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	74	52.9	697	4	BI826689	603077454
2	73.6	52.6	506	6	CB446475	700183 MA
3	73.6	52.6	624	6	CB446113	697902 MA
4	60.4	43.1	293	5	EX472815	DKFZDP686D
5	57.6	41.1	536	4	BI831357	603074343
6	56.6	40.4	999	9	CL464985	SAIL1231
7	54.4	38.9	299	9	CR163087	Reverse s
8	54.4	38.9	854	9	AG521132	Mus muscu
9	54.4	38.9	1391	2	BB455155	HVSMEM009
10	53.4	38.1	800	9	AG553158	Mus muscu
11	53.2	38.0	1102	9	CL463500	SAIL1187
12	52.4	37.4	872	5	EX385509	BX385509
13	52	37.1	887	8	AG192949	SP_1022 B
14	52	37.1	951	4	BI958037	HVSMEM001
15	52	37.1	1036	9	AG419622	Mus muscu
16	51.8	37.0	830	9	CNS028FQ	Tetraodon
17	51.8	37.0	898	9	AG381174	Mus muscu
18	51.6	36.9	1024	8	CC275787	CH261-133
19	51.4	36.7	744	9	AG490609	Mus muscu
20	51.4	36.7	866	9	AG520508	Mus muscu
21	51.2	36.6	743	2	BP631446	HVSMEM001
22	51.2	36.6	949	9	AG326267	Mus muscu
23	51.2	36.6	1072	7	CV520572	0089P0047
24	51.2	36.6	1229	8	CC319720	TAM32-14E

25	51.2	36.6	1560	9	CL081488	CH216-162
26	51.2	36.6	1686	8	CC189908	CH261-611
27	51.2	36.6	1971	8	CC207340	CH261-181
28	51	36.4	567	9	AG462381	Mus muscu
29	51	36.4	1031	9	AG324949	Mus muscu
30	51	36.4	1072	9	CL462217	SAIL115
31	50.8	36.3	425	7	CK089660	C024F74.3
32	50.8	36.3	773	8	BZ050814	jnf64ali.
33	50.8	36.3	909	9	CL462489	SAIL1166
34	50.6	36.1	876	9	AG395251	Mus muscu
35	50.6	36.1	904	9	AG049068	Pan trogl
36	50.6	36.1	933	9	CNS006XG	AL066047
37	50.6	36.1	1201	9	CNS016F9	Drosophil
38	50.4	36.0	591	9	AG124266	Pan trogl
39	50.4	36.0	668	9	AG139415	AG124266
40	50.4	36.0	713	8	BH944647	Pan trogl
41	50.4	36.0	723	9	AG123920	AG123920
42	50.4	36.0	770	9	AG038994	Pan trogl
43	50.4	36.0	998	9	AG388659	Mus muscu
44	50.4	36.0	1305	8	BZ694620	SP_Ba004
45	50.4	36.0	1308	8	CC229000	CH261-32N

ALIGNMENTS

BI826689 697 bp mRNA linear EST 04-OCT-2001  
603077454F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5169180 5',  
mRNA sequence.  
ACCESSION BI826689  
VERSION BI826689.1 GI:15938239  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 697)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-x@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM11420 row: 1 column: 13  
High quality sequence stop: 686.  
Location/Qualifiers  
1. 697  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5169180"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_119"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."

FEATURES  
source

ORIGIN

```

Query Match      52.9%; Score 74; DB 4; Length 697;
Best Local Similarity 98.8%; Pred. No. 1.4e-06;
Matches 85; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 55 GGGACGGGGCGGGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGG 114
    |||
Db 1 GGGACGGGGCGGGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACA-G 59

QY 115 GGAGGGGAGGTAATAGTGGGCCAG 140
    |||
Db 60 GGAGGGGAGGTAATAGTGGGCCAG 85

RESULT 2
LOCUS CB446475 506 bp mRNA linear EST 26-MAR-2003
DEFINITION 700183 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CB446475
VERSION CB446475.1 GI:29252857
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 506)
AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329
Plate: FOY8054 row: K column: 21
Seq primer: TAGAAGGCACATCGAGG.
Location/Qualifiers
source 1..506
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

FEATURES
source
Query Match 52.6%; Score 73.6; DB 6; Length 506;
Best Local Similarity 81.6%; Pred. No. 1.7e-06;
Matches 111; Conservative 0; Mismatches 19; Indels 6; Gaps 2;

QY 5 TCGGGAGGAGGATCCCGAAGCTCGGCGTTCGCGTCAGACGCCGGAGGGGACCGGGC 64
    |||
Db 359 TTGCGGGAGGGAACCTCCGAGCCTCGGCGTCAGCGTCAGACGTTGGGAGGGGACCGGGC 418

QY 65 GGGGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGGAGG 124
    |||
Db 419 GGGGAGCAG-AGGGGAGAAATGGGAGGACGAAGAGGAGG-----GAAGAGGGGAGGGAGG 472

QY 125 GTAAATAGTGGGCCAG 140
    |||
Db 473 GTAAATAGTGGGCTAG 488

RESULT 3
Query Match 52.6%; Score 73.6; DB 6; Length 506;
Best Local Similarity 81.6%; Pred. No. 1.7e-06;
Matches 111; Conservative 0; Mismatches 19; Indels 6; Gaps 2;

QY 5 TCGGGAGGAGGATCCCGAAGCTCGGCGTTCGCGTCAGACGCCGGAGGGGACCGGGC 64
    |||
Db 359 TTGCGGGAGGGAACCTCCGAGCCTCGGCGTCAGCGTCAGACGTTGGGAGGGGACCGGGC 418

QY 65 GGGGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGGAGG 124
    |||
Db 419 GGGGAGCAG-AGGGGAGAAATGGGAGGACGAAGAGGAGG-----GAAGAGGGGAGGGAGG 472

QY 125 GTAAATAGTGGGCCAG 140
    |||
Db 473 GTAAATAGTGGGCTAG 488

```

```

CB446113/c
LOCUS CB446113 624 bp mRNA linear EST 26-MAR-2003
DEFINITION 697902 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION CB446113
VERSION CB446113.1 GI:29252495
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 624)
AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329
Plate: FOY8054 row: K column: 21
Seq primer: GTAATAGGACTCACTATAGG.
Location/Qualifiers
source 1..624
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN
Query Match 52.6%; Score 73.6; DB 6; Length 624;
Best Local Similarity 81.6%; Pred. No. 1.7e-06;
Matches 111; Conservative 0; Mismatches 19; Indels 6; Gaps 2;

QY 5 TCGGGAGGAGGATCCCGAAGCTCGGCGTTCGCGTCAGACGCCGGAGGGGACCGGGC 64
    |||
Db 348 TTGCGGGAGGGAACCTCCGAGCCTCGGCGTCAGCGTCAGACGTTGGGAGGGGACCGGGC 289

QY 65 GGGGAGTAGTGGGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGGAGG 124
    |||
Db 288 GGGGAGCAG-AGGGGAGAAATGGGAGGACGAAGAGGAGG-----GAAGAGGGGAGGGAGG 235

QY 125 GTAAATAGTGGGCCAG 140
    |||
Db 234 GTAAATAGTGGGCTAG 219

RESULT 4
BX472815/c
LOCUS BX472815 293 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686D23150.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION DKFZp686D23150.5, mRNA sequence.
VERSION BX472815.1 GI:31667071
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 293)
AUTHORS Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amlid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.

```

TITLE EST (Bloeker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MIPS

Ingoletadter Landstr. 1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GfR (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.

No sl sequence available.  
This clone (DKFZp686D23150) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
1..293  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686D23150"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/notes="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

## ORIGIN

Query Match 43.1%; Score 60.4; DB 5; Length 293;  
Best Local Similarity 98.4%; Pred. No. 0.0014; Indels 0; Gaps 0;  
Matches 61; Conservative 0; Mismatches 1;  
QY 79 GAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGAGGTAATAGTGGCC 138  
|||||  
DB 293 GAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGAGGTAATAGTGGCC 234  
|||||  
QY 139 AG 140  
||  
DB 233 AG 232

RESULT 5  
BI831357  
LOCUS 536 bp mRNA linear EST 04-OCT-2001  
DEFINITION 603074343F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5166322 5',  
mRNA sequence.  
ACCESSION BI831357  
VERSION BI831357.1 GI:15942907  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 536)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11413 row: e column: 11  
High quality sequence stop: 536.

FEATURES  
source  
1..536  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clone="IMAGE:5166322"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_119"  
/notes="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."

ORIGIN  
Query Match 41.1%; Score 57.6; DB 4; Length 536;  
Best Local Similarity 93.8%; Pred. No. 0.0056;  
Matches 60; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 75 GCGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGTAATAGTG 134  
|||||  
DB 1 GCGGAGAAATGGGAGGACGAAGGGGAGGGAAGGACAGGGGAGGTAATAGTG 60  
|||||  
QY 135 GGCC 138  
|||  
DB 61 GGCC 64

RESULT 6  
CL464985/c

LOCUS  
DEFINITION SAIL\_1231\_H09.v1 SAIL Collection Arabidopsis thaliana genomic clone  
SAIL\_1231\_H09.v1, genomic survey sequence.

ACCESSION CL464985  
VERSION CL464985.1 GI:45867890  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,  
Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D.,  
Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmery, B.,  
Mittel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S. A.  
A high-throughput Arabidopsis reverse genetics system  
Plant Cell 14 (12), 2985-2994 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Sessions A  
Applied Trait Genetics  
Syngenta Biotechnology Inc.  
3054 Cornwalis Rd., Research Triangle Park, NC 27709, USA  
Email: allen.sessions@syngenta.com  
ABRC Stock Number CS945107; T-DNA left border flanking sequences of  
Syngenta Arabidopsis Insertion Library (SAIL) lines are available  
through the Arabidopsis Biological Resource Center (ABRC).  
Sequences represent a pool of amplified genomic regions and not  
single contiguous sequences.  
Class: TDNA tagged.

Location/Qualifiers

1..999  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="SAIL\_1231\_H09.v1"  
/clone\_lib="SAIL Collection"  
/note="T-DNA left border sequences were isolated using a  
modified TAIL-PCR strategy"

FEATURES  
source

1..999  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="SAIL\_1231\_H09.v1"  
/clone\_lib="SAIL Collection"  
/note="T-DNA left border sequences were isolated using a  
modified TAIL-PCR strategy"

## ORIGIN



**TITLE** Development of a genetically and physically anchored EST resource for barley genomics: Morex 5-45 DAP spike cDNA library

**JOURNAL** Unpublished (2001)

**COMMENT** On Jul 26, 2000 this sequence version replaced gi:9464547.

Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Total hg bases = 164  
Seq primer: AATTAAACCTCACTAAAGGG  
High quality sequence start: 248  
High quality sequence stop: 1317.

**FEATURES**

source

1. .1391  
Location/Qualifiers  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="Morex"  
/sub\_species="vulgare"  
/db\_xref="taxon:112509"  
/clone="HVSMEH0096G10f"  
/tissue\_type="5-45 DAP Spike"  
/lab\_host="SOLR"  
/clone\_lib="Hordeum vulgare 5-45 DAP spike EST library  
HVCdNA0009 (5 to 45 DAP)"  
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
plants were grown in the greenhouse at the University of  
California, Riverside (Penton, SJ Close, TJ Close). Whole  
spikes with awns trimmed were collected at 5, 10, 15, 20,  
30 and 45 DAP (Penton). Total RNA was prepared from each  
pool, equal quantities of all six RNA pools were combined,  
poly(A) RNA was purified from the mixture, one primary  
unamplified cDNA library was made, and 1 million pfu were  
in vivo excised to give pBluescript SK(-) cDNA phagemids  
(Choi) in the TJ Close lab at the University of California,  
Riverside. Phagemids were plated and picked at the Clemson  
University Genomics Institute (CUGI) (Begum, Palmer,  
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA  
sequencing and sequence analysis were performed at CUGI  
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).  
The sequence has been trimmed to remove vector sequence  
and contains a minimum of 100 bases of phred value 20 or  
above. For more details on library preparation and  
sequence analysis see  
http://www.genome.clemson.edu/projects/barley. To order  
this clone see http://www.genome.clemson.edu/orders/Also  
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

**ORIGIN**

Query Match 38.9%; Score 54.4; DB 2; Length 1391;  
Best Local Similarity 62.5%; Pred. No. 0.027;  
Matches 85; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGATCCCGAAGCTCGCGGTCTCGCGTCAGACGCCGGGAGGGGACG 60  
DB 817 GGG 876  
QY 61 GGGGGGGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGGGGGGAAGGACAGGGGAGGG 120  
DB 877 GGGGGGGGAGAGGG 936  
QY 121 GAGGGTAAATAGTGGG 136  
DB 937 TTGGGAATGGGTGGG 952

**RESULT 10**  
AG553158/c

**LOCUS** AG553158 800 bp DNA linear GSS 05-JUN-2004

**DEFINITION** Mus musculus molossinus DNA, clone:MSMg01-470B03.T7, genomic survey sequence.

**ACCESSION** AG553158

**VERSION** AG553158.1 GI:48313849

**KEYWORDS** GSS.

**SOURCE** Mus musculus molossinus

**ORGANISM** Mus musculus molossinus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE** 1 Hattori M., Toyoda A., Noguchi H., Kojima T. and Sakaki Y.  
BAC end Sequences of Library MSMg01

**AUTHORS** Hattori M., Toyoda A., Noguchi H., Kojima T. and Sakaki Y.

**TITLE** Direct Submission

**JOURNAL** Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan 1-7-22 Suehiro-chou, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

**COMMENT** Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Teukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@rtc.riken.jp  
PRIMERS  
Sequencing : T7  
LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI

**FEATURES**

source

1. .800  
Location/Qualifiers  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSMg01-470B03.T7"  
/sex="male"  
/tissue\_type="mixture of kidney and spleen"  
/clone\_lib="MSMg01 Mouse Male BAC Library"

**ORIGIN**

Query Match 38.1%; Score 53.4; DB 9; Length 800;  
Best Local Similarity 62.2%; Pred. No. 0.046;  
Matches 84; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1 GGGTTCGGGAGGAGATCCCGAAGCTCGCGGTCTCGCGTCAGACGCCGGGAGGGGACG 60  
DB 712 GGGGGCTGTAGGGGGGAGGGGGGGGGGAGGGGGGAGGGAGGGGGGAGAGGGAG 653  
QY 61 GGGGGGGGGAGTGTAGTGGGGGAGAAATGGGAGGACGAAGGGGGGAGGAGGAGGGGAGGG 120  
DB 652 GGG 593  
QY 121 GAGGGTAAATAGTGGG 135  
DB 592 GAGGGGGGTATAGGGG 578

**RESULT 11**  
CL463500

**LOCUS** CL463500 1102 bp DNA linear GSS 31-MAR-2004

**DEFINITION** SAIL1187 All.v1 SAIL Collection Arabidopsis thaliana genomic clone

**ACCESSION** CL463500

**VERSION** CL463500.1 GI:45866405

**KEYWORDS** GSS.

**SOURCE** Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 1102)

AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McEliver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmery,B., Mittel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.  
A high-throughput Arabidopsis reverse genetics system  
Plant Cell 14 (12), 2985-2994 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Applied Trait Genetics  
Syngenta Biotechnology Inc.  
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
Email: allen.sessions@syngenta.com  
ABRC Stock Number CS943930; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).  
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.  
Class: TDNA tagged.  
Location/Qualifiers  
1. .1102  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="SAIL 1187 A11.v1"  
/clone\_lib="SAIL Collection"  
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN  
Query Match 38.0%; Score 53.2; DB 9; Length 1102;  
Best Local Similarity 63.2%; Pred. No. 0.05;  
Matches 79; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 GGTTTCGGAGGAGGATCCGAGGCTCGCGGTGTCGCTCAGACCCCGGGAGGGGAGG 60  
|||||  
Db 740 GGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 799  
|||||

QY 61 GGGCGGGAGTACTGGGGAGGATCGGAGGACCAAGGGGAGGAGGAGGAGGAGGAGG 120  
|||||  
Db 800 GGGAGGGGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 859  
|||||

QY 121 GAGGG 125  
|||||  
Db 860 GAGGG 864  
|||||

RESULT 12  
BX385509  
LOCUS  
DEFINITION  
BX385509 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
Homo sapiens cDNA clone CS0DL012YC09 3-PRIME, mRNA sequence.

ACCESSION  
BX385509  
VERSION  
BX385509.2 GI:46622615  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE 1 (bases 1 to 872)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 8, 2003 this sequence version replaced gi:30455339.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 1102)

AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McEliver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmery,B., Mittel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.  
A high-throughput Arabidopsis reverse genetics system  
Plant Cell 14 (12), 2985-2994 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Applied Trait Genetics  
Syngenta Biotechnology Inc.  
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
Email: allen.sessions@syngenta.com  
ABRC Stock Number CS943930; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).  
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.  
Class: TDNA tagged.  
Location/Qualifiers  
1. .1102  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="SAIL 1187 A11.v1"  
/clone\_lib="SAIL Collection"  
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN  
Query Match 38.0%; Score 53.2; DB 9; Length 1102;  
Best Local Similarity 63.2%; Pred. No. 0.05;  
Matches 79; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 GGTTTCGGAGGAGGATCCGAGGCTCGCGGTGTCGCTCAGACCCCGGGAGGGGAGG 60  
|||||  
Db 740 GGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 799  
|||||

QY 61 GGGCGGGAGTACTGGGGAGGATCGGAGGACCAAGGGGAGGAGGAGGAGGAGGAGG 120  
|||||  
Db 800 GGGAGGGGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 859  
|||||

QY 121 GAGGG 125  
|||||  
Db 860 GAGGG 864  
|||||

RESULT 12  
BX385509  
LOCUS  
DEFINITION  
BX385509 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
Homo sapiens cDNA clone CS0DL012YC09 3-PRIME, mRNA sequence.

ACCESSION  
BX385509  
VERSION  
BX385509.2 GI:46622615  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE 1 (bases 1 to 872)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 8, 2003 this sequence version replaced gi:30455339.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6730.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DL012AB05NP1&c=6730.r.

FEATURES  
source  
1. .872  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DL012YC09"  
/cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
/cell\_line="RAMOS CELL LINE"  
/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the PCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 37.4%; Score 52.4; DB 5; Length 872;  
Best Local Similarity 56.6%; Pred. No. 0.076;  
Matches 77; Conservative 9; Mismatches 50; Indels 0; Gaps 0;

QY 1 GGTTTCGGAGGAGGATCCGAGGCTCGCGGTGTCGCTCAGACCCCGGGAGGGGAGG 60  
|||||  
Db 378 GSGAGGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 437  
|||||

QY 61 GGGCGGGAGTACTGGGGAGGATCGGAGGACCAAGGGGAGGAGGAGGAGGAGGAGG 120  
|||||  
Db 438 GGGCGGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 497  
|||||

QY 121 GAGGGTAAATAGTGGG 136  
|||||  
Db 498 AAGRRCAAGNAGGG 513  
|||||

RESULT 13  
AZ192949  
LOCUS  
DEFINITION  
SP\_1022\_B1\_C01\_T7A Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=1022 Col=1 Row=F, genomic survey sequence.

ACCESSION  
AZ192949  
VERSION  
AZ192949.1 GI:8376128  
KEYWORDS  
GSS.  
SOURCE  
Strongylocentrotus purpuratus  
ORGANISM  
Strongylocentrotus purpuratus

REFERENCE 1 (bases 1 to 887)

AUTHORS Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R., Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,B.H. and Hood,L.  
A sea urchin genome project: Sequence scan, virtual map, and additional resources  
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: Cameron, RA, Davidson, EH, Hood, L  
Division of Biology 156-29  
California Institute of Technology  
Pasadena California 91125, USA  
Tel: (626) 395-8421  
Fax: (626) 793-3047  
Email: acameron@caltech.edu

FEATURES	source
high qual	1
Location/Qualifiers	951
organism="Hordeum vulgare subsp. vulgare"	
/mol_type="mRNA"	
/cultivar="Morex"	
/sub_species="vulgare"	
/db_xref="taxon:112509"	
/clone="HVMENC011B04f"	

